THESIS

DESIGN AND QUANTIFICATION OF A TISSUE TYPE SPECIFIC GENETIC CIRCUIT IN PLANTS

Submitted by

Sara Oehmke

Graduate Degree Program in Cell and Molecular Biology

In partial fulfillment of the requirements

For Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2020

Master's Committee:

Advisor: June Medford

Cris Argueso
Marc Nishimura

Copyright by Sara Elizabeth Oehmke 2020

All Rights Reserved

ABSTRACT

DESIGN AND QUANTIFICATION OF A TISSUE TYPE SPECIFIC GENETIC CIRCUIT IN PLANTS

Synthetic biologists aim to rationally design genetic circuits and utilize plant platforms to photosynthetically drive, self-sustainable circuits. Although plants are excellent platforms, issues and unpredictability arise from the innate complexity of multicellularity.

The ability to quantitatively control gene expression within specific cell types can address some issues arising from multicellularity. In my research, I developed a genetic circuit with the ability to induce and quantitatively control output of a genetic circuit in *Arabidopsis thaliana* root epidermal cells. The circuit design uses an externally applied ligand that activates a computationally designed transcriptional response driven by a tissue specific promoter to control output (GFP expression). In addition, I engineered a circuit that adds a positive feedback motif. To quantify the behaviors of these circuits I developed a MATLAB program to remove background signals from confocal microscopy images and quantify GFP signal in a high-throughput manner.

The genetic circuit is highly specific for root epidermal cells, controllable with external ligand, and has increased sensitivity and memory with positive feedback. The concepts and components of these circuits can

be implemented in future designs to engineer and produce plants with more predictable and diverse behaviors affording the operator greater control.

TABLE OF CONTENTS

Chapter 1: Introduction

1.1 Synthetic biology

Synthetic biology is an interdisciplinary field incorporating engineering and computational principles and approaches to rationally (re)design natural and *de novo* biological systems [1]. By invoking forward engineering, the design-test-build cycle, and *in silico* methods (mathematical modeling), synthetic biologists have rationally engineered biological systems. These systems perform *desired* behaviors (functions) laid out in designs in a highly predictable and robust manner [2] [3]. This contrasts with earlier bioengineering efforts that produced biological systems with unpredictable behaviors. A key to designing predictable and robust behavior is applying transfer functions. Transfer functions quantify how a control variable (input) within a biological system can produce a response (output) [4].

Transfer functions, also known as input-output functions, are used to develop computational models to predict behaviors of biological systems *in silico*. These models are informed by empirical data to adjust parameters within the model and/or system designs. This is an iterative process wherein feedback between models and experimental data is utilized to tune responses and outputs [5]. Recently, a library of tuned inducible synthetic bacterial promoters was generated using feedback between a thermodynamic model and data derived from cell-based assays [6]. The model described the affinity of transcriptional machinery to promoters. It

predicted that varying specific sequences within the promoter could adjust the affinity of transcriptional machinery to the promoter, and thereby change transcription rates. By modifying the transcription rate, the dynamic range, *i.e.*, the absolute difference between the “ON” and “OFF” states, could be tuned. The result of several iterations of feedback between the thermodynamic model and experimental data was a library of inducible systems with predictable dynamic ranges. Furthermore, four varied regulatory sequences from the library were used to design multi-input Boolean logic gate circuits with highly predictable outputs. Synthetic biologists have invoked this iterative process to design many other synthetic systems which function as switches, oscillators, and Boolean logic gates [7] [8] [9].

It would be extremely advantageous to apply synthetic biology approaches to design and redesign biological systems that function within plants. First, plants are self-sustainable and can power synthetic circuity autonomously via photosynthesis. Second, there are many applications for rationally designed plants. As suggested, plants could be designed to grow as structures, have fruits with juice sacs containing medicines, or grow to a specific developmental stage and switch to biofuel production [4] [10] [11]. Additionally, agricultural crops can be engineered and optimized. Crops have been designed to be more resistant to abiotic and biotic challenges, be more nutritional, and utilize resources more efficiently [12] [13] [14] [15]. Applying synthetic biology approaches could refine the process of engineering plants,

cutting down the time required to tune designs, and create plants with more robust behavior [9]. Furthermore, plants are a major source of food, fiber, and renewable materials around the world, and required for life on Earth, the ability to engineer and design plants could change the world.

1.2 A brief overview of synthetic biology

The first wave of synthetic biology began at the end of the last century, and focused on rational design of synthetic devices and modules implemented in prokaryotic platforms, namely *Escherichia coli*. During this period seminal works such as oscillators, toggle switches, and various types of Boolean logic gates were engineered using mathematical modeling. These works functioned as proof-of-concept that rational design with mathematical modeling could result in genetic systems with more predictable outputs [7] [8].

The second wave of synthetic biology aims to scale up smaller genetic devices and modules into fully programmable systems, and possibly engineer whole cells [5]. An important aspect of the second wave is compiling large libraries of specific genetic parts, *e.g.*, BioBricks [16]. The large number of available and quantified parts which can perform redundant functions strives to maintain gene parts or modules that act independently of endogenous systems and each other within higher order systems, *i.e.* orthogonality [17][18]. Orthogonality within a biological system permits more predictable behavior by eliminating cross-talk. Furthermore, orthogonality permits interoperability between individual genetic circuits and modules to

form programmable systems. These systems are increasingly more complex as the number of quantitatively and functionally characterized modules and parts increases. Individual cells have been genetically programmed to perform coordinated tasks with other genetically modified cells within populations. A few examples of coordinated behavior include pattern formation, biofilm aggregation, and even self-organization of simple synthetic tissue [19] [20] [21] [22].

The foundation of synthetic biology was built on applying engineering and mathematical modeling approaches to design synthetic biological systems which function within bacteria [23]. Fundamentals and principles applied during this earlier period are being invoked to engineer synthetic systems which function in eukaryotic platforms in a goal-oriented manner [11][24]. These applications are what make eukaryotes such an attractive host for synthetic systems. More specifically, engineered plants can be applied to human health, resource management, and energy and environmental issues in manners unicellular organisms simply cannot [25]. However, multicellularity presents many challenges when designing synthetic systems.

1.3 Control of gene expression within multicellular organisms

Multicellularity is an attractive trait to exploit but challenging because gene expression is tightly controlled and regulated within multicellular organisms. The layers of internal regulation coupled with external input underlying emergent behavior, such as cell differentiation, can modify

synthetic circuit behaviors [26] [27]. The architecture and dynamics of eukaryotic DNA are far more complex than bacterial DNA architecture and dynamics [28]. The chromatin state of DNA can silence or activate genes, and produce contextual effects on heterologous DNA, particularly if insertion events are random [25]. “Landing pads” have been developed to prevent confounding positional effects by controlling what loci heterologous DNA is inserted into [29]. However, these systems have yet to be developed for many eukaryotic hosts. Additionally, regulatory elements can be located hundreds to thousands of base pairs away from concomitant sequences under regulation. Interactions between transcriptional machinery and far away regulatory elements can alter DNA architecture, and modify gene expression further [28]. Moreover, the distance in itself can make identifying regulatory elements difficult. Although, bacterial chromosomes also have upstream and downstream regulatory elements affecting DNA architecture the distance between these elements and genes is smaller compared with eukaryotic chromosomes [28]. Additionally, eukaryotic cells perform post-translational modifications, alternative-splicing, and polyadenylation [25] [26]. All of this genetic regulation can affect dynamics, kinetics, and behaviors of synthetic genetic circuits and systems.

Constitutive expression of heterologous circuitry may also present issues when utilizing multicellular platforms for synthetic systems. Constitutive expression is defined as gene expression regardless of tissue and cell type, or developmental and temporal regulation. It can disrupt critical

developmental trajectories, metabolic networks, and become toxic and/or fatal [30]. Spatially targeted and temporally controlled gene expression may be desirable depending on the designed response or downstream application of the synthetic system. One way to address the challenges and issues arising from multicellularity and constitutive expression is to produce a tissue specific genetic controller, which quantitatively controls gene expression by an external input.

Temporal and quantitative control of gene expression can be genetically encoded by inducible genetic systems. Inducible systems can control when gene expression occurs. They function to turn a genetic circuit “ON” or “OFF” at a given developmental stage, under specific contexts, or within certain time frames [5]. Moreover, gene expression can be quantitatively controlled with inducible systems by adjusting the concentration of the externally inputted molecule, *e.g.,* ligands, eliciting the inducible response. [31]. Furthermore, proteins, *e.g.,* transcription factors, can be computationally designed such that they interact with specific pertinent ligands and can activate transcription after exposure to the ligand [31,32,33]. This computational design approach could be leveraged to develop a variety of inducible systems acting as biosensors.

Cell and tissue specific regulatory elements, *e.g*., promoters, can target expression of regulatory biomolecules, *e.g.*, transcription factors, and RNA, *i.e., to* spatially control gene expression [34]. For example, an anther specific promoter driving the expression of a cytotoxic gene ablated the anther cell

line, producing a male sterility phenotype in *Brassica napus* plants [35]. Constitutive expression of the cytotoxic gene would have been fatal. However, targeted gene expression rendered the desired male-sterility phenotype.

Inducible and tissue-specific systems are utilized in functional analysis of gene expression, particularly in developmental studies, and connecting genotype to phenotype. For example, trapper *Arabidopsis thaliana* lines with Gal-4 inducible systems and tissue specific promoters have been utilized to observe and characterize expression patterns [34]. This can be a valuable asset for designing synthetic genetic circuitry. First, genetic circuits can be redesigned with targeted cell-line or tissue specific responses under quantitative control to leverage multicellularity. Second, synthetic components can be functionally and quantitatively characterized leading to development of higher order genetic systems implemented in plant hosts.

1.4 Regulatory motifs

DNA regulatory motifs are utilized in synthetic circuits to genetically encode dynamic responses. Synthetic genetic circuitry imparts the ability to sense signals, compute, and actuate a desired response or behavior [18]. Some examples of dynamic responses include switches and cellular "memory," amplification, and oscillation [7] [8] [36]. Various circuits with these dynamic responses are designed to exclusively utilize repression, or activation, or a combination of the two. Repression and activation occur as a result of interactions between DNA regulatory elements, *e.g.*, promoters, and

biomolecules, *e.g.*, transcription factors. It is therefore critical that genetic circuits contain specific parts and modules to build regulatory motifs.

To achieve a desired response the appropriate genetic parts and modules are required [17]. The specificity, affinity, and interactions between biomolecules and DNA can determine the dynamic responses of a genetic circuit. Furthermore, qualitative functions of circuits also depend on how these parts and modules are wired together, *i.e.*, the topology of circuits [18] [23]. For instance, a toggle switch can be modified into a clock by changing the relative expression of the transcriptional repressors that act to repress each other within the circuit [38]. It is therefore critical that the parts and modules used, and how they are arranged is rationally designed to program cells with the desired behavior.

1.4.2 Feedback and biological switches

Genetic circuits comprised of appropriate regulatory elements, and repressors or activators can be arranged into feedback systems. Feedback acts either to amplify the product of a reaction (positive feedback), or to diminish the product of a reaction (negative feedback). Feedback motifs are critical components of biological switches [38] [39].

Biological switches can be defined as regulatory systems comprised of genetic elements or proteins or a combination of these controlling the state of a cell. Biological switches are ubiquitous in nature and critical to many cellular processes including cell differentiation, cell-cycle progression, and

apoptosis [40] [41]. Simple synthetic genetic biological switches have been engineered to toggle between cellular states by modifying the level of a molecule within a cell between high and low concentrations [7].

Both natural and synthetic switches must have at least one "ultrasensitive" component [42][43]. Ultrasensitivity is a biological phenomenon wherein a small change of input stimulus is able to drive output to a much greater percentage of the maximum output [42][43]. One possible consequence of ultrasensitivity is a graded output, which produces an amplification response [44]. Alternatively, ultrasensitivity can produce an "ON"/"OFF" switch. This is visually represented as a sharp sigmoidal input-output curve with low and high steady-states. This is known as bistability [23]. Bistable genetic circuits with switch-like behavior can impart cellular "memory," or hysteresis through transcription [36]. This behavior rises from the unstable state that is in between the low and high steady-states acting as a switch.

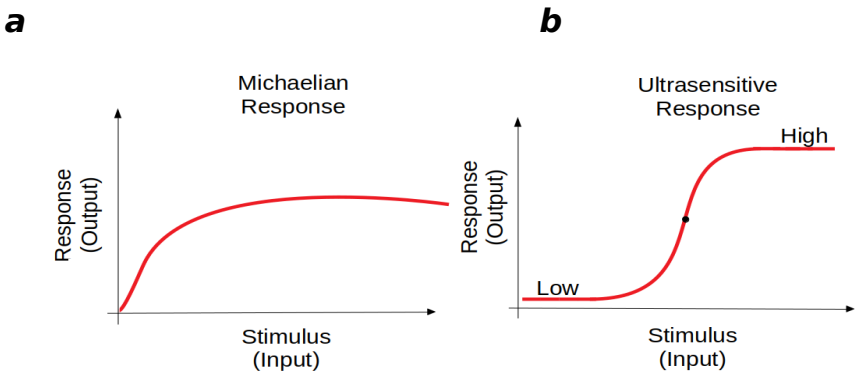


Figure 1.1. Input-output graphs of Michaelian and Ultrasensitive responses. Graph *a* depicts a Michaelian response which is the simplest mass action system. The output is graded as input increases until the system is saturates and plateaus. Graph *b* depicts an ultrasensitive response

wherein two stable steady states exist. A low steady state, at low input, and a high steady state at high input. In-between these steady states, represented by the dot, is an unstable state acting as the switch. Notice the sharp nearly linear increase between low and high steady states.

Memory can be defined as a sustained response to a single transient stimulus. That is, the current state of a cell is dependent on its molecular history. Bistable synthetic circuits with memory have been implemented in bacterial and eukaryotic hosts [36][37]. For example, Ajo-Franklin et al., developed a genetic circuit with an autoregulatory positive feedback, and observed bistable and memory responses within yeast. In this study, activator-promoter pairs were quantitatively characterized to produce a model to inform memory system design. While empirical data showed growth rate determined if the circuit was bistable with memory responses (high growth rate) or monostable without memory (low growth rate). By factoring the growth rate into the model, the behavior of the circuit could be tuned.

1.5 Plant confocal microscopy and the autofluorescence issue

Quantification of genetic part and module behaviors is a major bottleneck when developing synthetic genetic circuits implemented in plant platforms [9][25]. A strategy to observe genetic circuit and biomolecule dynamics, and quantify behaviors is through fluorescent microscopy. Specifically, observation can be experimentally achieved by fusing fluorescent molecules to a biomolecule, or by driving expression of

fluorescent molecules with pertinent promoters. In this manner, fluorescent molecules act as reporters, and the intensity of the signal, or emitted light from the reporters can be quantified within images, and thereby the dynamics and behaviors of a genetic circuit and parts of those circuit can be quantified. However, an issue arises from exciting fluorescent reporters because living tissues, particularly plant tissues, also naturally produce biomolecules and biopolymers, such as NADH, chlorophyll, cell wall polymers and components, and pyridine that emit fluorescence when excited by the same wavelengths of light used to excite fluorescent reporters [45]. Collectively, this is known as autofluorescence. Autofluorescence can partially occlude and artificially inflate reporter fluorescence and signal. If a genetic module or circuit had a low level of output or expressed reporter at a low level such that autofluorescence signal intensity was greater than the fluorescence signal, the reporter or output signal would not be discernible from autofluorescence [46]. Results could be interpenetrated incorrectly, specifically as the absence of reporter or output. Convexly, reporter or output signal with a greater intensity than autofluorescence would be inflated by autofluorescence, consequently measurements of report or output signal would be inflated. Either influence of autofluorescence can lead to inaccurate quantification and faulty interpretation of genetic circuit dynamics. To accurately quantify reporter or genetic circuit output by fluorescent intensity, autofluorescence must be removed from fluorescent microscopic images.

1.6 Thesis statement and overview

1.6.1 Thesis statement

This thesis demonstrates approaches to design, and qualify and quantify behavior of an inducible, root epidermal cell-specific genetic circuit. Additionally, I have detailed the changes in the dynamics and kinetics of this genetic circuit with the addition of a positive feedback motif.

1.6.2 Thesis overview

Following the introductory chapter, chapter 2 details the design of and experimental approaches to build, qualify, and quantify an inducible, root epidermal cell specific expression genetic circuit. This chapter focuses on fluorescent confocal microscopy and computational approaches to remove autofluorescence and background to accurately quantify gene expression. Chapter 3 discusses the design of a positive feedback loop, and incorporation of this loop into the root epidermal cell-specific genetic circuit. Additionally, qualification and quantification of gene expression after incorporation of a positive feedback loop. Furthermore, this chapter details comparisons between the inducible root epidermal cell-specific genetic circuit with and without the positive feedback regulatory motif.

REFERENCES

[1] Ahmad. K.J., & Collins, J.J. “Synthetic Biology: Applications Come of Age.” *Nature Reviews Genetics*, vol. 11, no. 5, May 2010, pp. 367–379., doi:10.1038/nrg2775.

[2] Wang B., Kitney R. I., Joly N., & Buck M. “Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology.” *Nature Communications,* vol. 2, no. 58, 2011. 10.1038/ncomms1516

[3] Daniel, R., Rubens, J., Rahul S., & Lu T. “Synthetic Analog Computation in Living Cells.” *Nature*, vol. 497, no. 7451, Nature Publishing Group, 2013, pp. 619–23, doi:10.1038/nature12148.

[4] Medford, J.I., & Prasad, Ashok “Towards Programmable Plant Genetic Circuits.” *The Plant Journal : For Cell and Molecular Biology*, vol. 87, no. 1, 2016, pp. 139–48, doi:10.1111/tpj.13235.

[5] Slusarczyk, A.L., Lin, A., & Weiss, R. “Foundations for the design and implementation of synthetic genetic circuits.” *Nature*, vol. 13, Macmillan Publishers, 2012, doi:10.1038/nrg/3227.

[6] Chen, Y., Ho, J., Shis,D., Gupta, C., Long, J., Wagner, D., Ott, W., Josic, K. & Bennett, M. “Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors.” *Nature Communications*, vol. 9, no. 64, doi.org/10.1038/s41467-017-02473-5.

[7] Gardner, T., Cantor, C., & Collins, J.J. “Construction of a Genetic Toggle Switch in Escherichia Coli.” *Nature*, vol. 403, no. 6767, 2000, pp. 339–42, doi:10.1038/35002131.

[8] Elowitz, M.B., & Leibler, S. “A Synthetic Oscillatory Network Repressilator.” *Nature*, vol. 403, 2000, pp. 335–38, doi:10.1038/35002125.

[9] Schaumberg, K.A., Antunes, M.S., Kassaw, T.K., Xu, W., Zalewski, C.S., Medford, J.I., & Prasad, A. “Quantitative Characterization of Genetic Parts and Circuits for Plant Synthetic Biology.” *Nature Methods*, vol. 13, no. 1, Nature Publishing Group, Dec. 2015, pp. 94–100, doi:10.1038/nmeth.3659

[10] Medford, J.I., & Prasad, A. “Plant Synthetic Biology Takes Root.” *Science*, vol. 346, no. 6206, Sept. 2014, pp. 162–163., doi:10.1126/science.1261140.

[11] Medford, J. I., & McCarthy, D. M. (2017). “Growing beyond: Designing Plants to Serve Human and Environmental Interests.” *Current Opinion in Systems Biology*, vol. 5, 2017, pp. 82–85., doi:10.1016/j.coisb.2017.08.008.

[12] Paddon, C.J., Westfall, P.J., & Newman, J.D., Pitera, D.J. Benjamin K., Fisher, K., McPhee, D., Tai, A., Main, A., Eng, D., Polichuk, D.R., Teoh, K.H., Reed, D.W., Treynor, T., Lenihan, J., Jiang, H., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K.W., Fickes, S., Galazzo, J., Gaucher, S.P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., L.F. Westblade, Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P.S., Keasling, J.D., Reiling, K.K., Renninger, N.S., & Newman, J.D. “High-Level Semi-Synthetic Production of the Potent Antimalarial Artemisinin.” *Nature*, vol. 496, no. 7446, 2013, pp. 528–32, doi:10.1038/nature12051.

[13] Dalal, J., Lopez,H., Vasani, N., Hu, Z., Swift, J.E., Yalamanch, R., Dvora, M., Lin, X., Xie, D., Qu, R., & Sederoff, H. “A photorespiratory bypass increases plant growth and seed yield in biofuel crop *Camelina sativa*.” *Biotechnology for Biofuels*, vol. 8, no. 175, 2015, doi.org/10.1186/s13068-015-0357-1]

[14] Tay, P.K.R., Nguyen, P.Q., & Joshi, N.S. “A Synthetic Circuit for Mercury Bioremediation Using Self-Assembling Functional Amyloids.” *ACS Synthetic Biology*, vol. 6, no. 10, 2017, pp. 1841–50, doi:10.1021/acssynbio.7b00137.

[15] Siu, Y., Fenno, J., Lindle, J.M., & Dunlop, M.J. “Design and Selection of a Synthetic Feedback Loop for Optimizing Biofuel Tolerance.” *ACS Synthetic Biology*, vol. 7, no. 1, 2018, pp. 16–23, doi:10.1021/acssynbio.7b00260.

[16] Knight, T., Rettberg, R., Chan, L., Endy, D., Shetty, R., & Che, A. "Idempotent vector design for standard assembly of BioBricks," BBF RFC 9, 2003.

[17] Brophy, J. A. N. & Voigt, C. A. "Principles of genetic circuit design." *Nature Methods* vol. 11, no. 6, 2014, pp. 508–520.

[18] Purnick, P.E. M., & Weiss, R. “The Second Wave of Synthetic Biology: From Modules to Systems.” *Nature Reviews Molecular Cell Biology*, vol. 10, no. 6, Nature Publishing Group, 2009, pp. 410–22, doi:10.1038/nrm2698.

[19] Basu, S., Gerchman, Y., Collins, C.H., Arnold, F.H., & Weiss, R. “A synthetic multicellular system for programmed pattern formation.” *Nature,* vol 434, 2005, pp. 1130–1134.

[20] Danino, T., Mondragon, O., Tsimring, L., & Hasty, J. “A Synchronized Quorum of Genetic Clocks.” *Nature*, vol. 463, no. 7279, 2010, pp. 326-30, doi:10.1038/nature08753

[21] Hong, S.H., Hedge, M., Kim, J., Wang, X., Jayaraman, A., & Wood, T. “Synthetic Quorum-Sensing Circuit to Control Consortial Biofilm Formation and Dispersal in a Microfluidic Device.” *Nature Communications*, vol. 3, no. 1, 2012, doi:10.1038/ncomms1616.

[22] Johnson, M.B., March, A., & Morsut, L. “Engineering Multicellular Systems: Using Synthetic Biology to Control Tissue Self-Organization.” *Current Opinion in Biomedical Engineering*, vol. 4, 2017, pp. 163–73, doi:10.1016/j.cobme.2017.10.008.

[23] Cameron, D.E., Bashor, C.J., & Collins, J.J. “A brief history of synthetic biology.” *Nature Reviews Microbiology*, vol. 12, 2014, pp. 381-390. doi:https://doi.org/10.1038/nrmicro3239.

[24] Antunes, M.S., Morey, K.J., Smith, J.J., Albrecht, K.D., Bowen, T.A., Zdunek, J.K., Troupe, J.F., Cuneo, M.J., Webb, C.T., Hellinga, H.W., & Medford, J.I. “Programmable Ligand Detection System in Plants through a Synthetic Signal Transduction Pathway.” *PLoS ONE*, vol. 6, no. 1, 2011, doi:10.1371/journal.pone.0016292.

[25] Ceroni, F., & Ellis. T. “The Challenges Facing Synthetic Biology in Eukaryotes.” *Nature Reviews Molecular Cell Biology*, vol. 19, no. 8, 2018, pp. 481–482., doi:10.1038/s41580-018-0013-2.

[26] James, A.B., Syed, N.H., Bordage, S., Marshall, J., Nimmo, G.A., Jenkins, G.I., Herzyk, P., Brown, J.W.S., & Nimmo, H.G. “Alternative Splicing Mediates Responses of the Arabidopsis Circadian Clock to Temperature Changes.” *Plant Cell*, vol. 24, no. 3, 2012, pp. 961–81, doi:10.1105/tpc.111.093948.

[27] Ptitsyna, N., Boughorbel, S., Anbari, M.E., & Ptitsyn, A. “The Role of Alternative Polyadenylation in Regulation of Rhythmic Gene Expression.” *BMC Genomics*, vol. 18, no. 1, BMC Genomics, 2017, pp. 1–8, doi:10.1186/s12864-017-3958-1.

[28] Garcia, H.G., Grayson, P., Han, L., Inamdar, M., Kondev, J., Nelson, P.C., Phillips, R., Widom, J., & Wiggins, P.A. “Biological Consequences of Tightly Bent DNA: The Other Life of a Macromolecular Celebrity.” *Biopolymers*, vol. 85, no. 2, 2007, pp. 115–130., doi:10.1002/bip.20627.

[29] Gaidukov, L., Wroblewska, L., Teague, B., Nelson, T., Zhang, X., Liu, Y., Jagtap, K., Mamo, S., Tseng, W.A., Lowe, A., & Das, J. “A Multi-Landing Pad

DNA Integration Platform for Mammalian Cell Engineering.” *Nucleic Acids Research*, vol. 46, no. 8, 2018, pp. 4072–86, doi:10.1093/nar/gky216.

[30] Koo, B.C, Kwon, M.S., Kim, & D., Kim, S.A. “Production of Transgenic Chickens Constitutively Expressing Human Erythropoietin (HEPO): Problems with Uncontrollable Overexpression of HEPO Gene.” *Biotechnology and Bioprocess Engineering*, vol. 22, no. 1, 2017, pp. 22–29., doi:10.1007/s12257-016-0590-x.

[31] Feng, J., Jester, B.W., Tinberg, C.E., Mandell, D.J., Antunes, M.S., Chari, R., Morey, K.J., Rios, X., Medford, J.I., Church, G.M., Fields, S., & Baker, D. “A General Strategy to Construct Small Molecule Biosensors in Eukaryotes.” *ELife*, vol. 4, 29 Dec. 2015, doi:10.7554/eLife.10606.

[32] Tinberg, C.E., Khare, S.D., Dou, J., Doyle, L., Nelson, J.W., Schena, A., Jankowski, W., Kalodimos, C.G., Johnsson, K., Stoddard, B.L., & Baker, D. “Computational Design of Ligand-Binding Proteins with High Affinity and Selectivity.” *Nature*, vol. 501, no. 7466, Sept. 2013, pp. 212–216., doi:10.1038/nature12443.

[33] Bick, M.J., Greisen, P.J., Morey, K.J., Antunes, M.S., La, D., Sankaran, B., Reymond, L., Johnsson, K., Medfor, J.I., & Baker. D. “Computational Design of Environmental Sensors for the Potent Opioid Fentanyl.” *ELife*, vol. 6, 19 Sept. 2017, doi:10.7554/elife.28909.

[34] Radoeva, T., Hove, C.A.T, Saiga, S., & Weijers, D. “Molecular Characterization of Arabidopsis GAL4/UAS Enhancer Trap Lines Identifies Novel Cell Type-Specific Promoters.” Plant Physiology, 2016, doi:10.1104/pp.16.00213.

[35] Roque, E., Gómez, M.D, Ellul, P., Wallbraun, M., Madueño, F., Beltrán, J.P., & Cañas, L.A. “The PsEND1 Promoter: A Novel Tool to Produce Genetically Engineered Male-Sterile Plants by Early Anther Ablation.” *Plant Cell Reports*, vol. 26, no. 3, 2007, pp. 313–25, doi:10.1007/s00299-006-0237-z.

[36] Ajo-franklin, C.M., Drubin, D.A., Eskin, J.A., Gee, E.P.s., Landgraf, D., Phillips, I., & Silver, P.A. “Rational Design of Memory in Eukaryotic Cells.” *Genes Dev*.-2007-Ajo-Franklin-2271-6. no. 617, 2007, pp. 2271–76, doi:10.1101/gad.1586107.GENES.

[37] Zedao, L., Zhang, J., Jin, J., Geng, Z., Qi, Q., & Liang, Q. “Programming Bacteria With Light—Sensors and Applications in Synthetic Biology.” *Frontiers in Microbiology*, 2018, vol. 9. doi: 10.3389/fmicb.2018.02692.

[38] Ellis T., Wang X., &Collins J.J. "Diversity-based, model-guided construction of synthetic gene networks with predicted funcitons. *Nature Biotech*., vol.27, pp.465-471.

[39] Lai, K., Robertson, M.A., & Schaffer, D.V. “The Sonic Hedgehog Signaling System as a Bistable Genetic Switch.” *Biophysical Journal,* vol. 86, no. 5, Elsevier, 2004, pp. 2748–57, doi:10.1016/S0006-3495(04)74328-3.

[40] Matten, W.T., Copeland, T.D., Ahn, N.G., &, Woude, G.F.V. “Positive Feedback between MAP Kinase and Mos DuringXenopusOocyte Maturation.” *Developmental Biology*, vol. 179, no. 2, 1996, pp. 485–492., doi:10.1006/dbio.1996.0277.

[41] Harris, Sandra L., & Arnold J. Levine. “The P53 Pathway: Positive and Negative Feedback Loops.” Oncogene, vol. 24, no. 17, 2005, pp. 2899–908, doi:10.1038/sj.onc.1208615.

[42] Goldbeter A., & Koshland, D.E. An amplified sensitivity arising from covalent modification in biological systems. PNAS. 1981; 78: 6840-6844

[43] Goldbeter, A., & Koshland, D.E. “Ultrasensitivity in Biochemical Systems Controlled by Covalent Modification.” The Journal of Biological Chemistry, vol. 259, no. 23, 1984, pp. 14441–47.

[44] Koshland D.E., Goldbeter, A., & Stock, J.B. Amplification and adaptation in regulatory and sensory systems. Science. 1982; 217:220–225.

[45] García-Plazaola, J.I, Fernández-Marín, B., Duke, S.O., Hernández, A., López-Arbeloa, F., &, Becerril, J.M. “Autofluorescence: Biological Functions and Technical Applications.” Plant Science, vol. 236, 2015, pp. 136–45, doi:10.1016/j.plantsci.2015.03.010.

[46] Billington, N., & Knight, A.W. “Seeing the Wood Through the Trees. Analytical Biochemistry, vol. 291, 2001, p. 175-197.

Chapter 2: Root epidermal cell-specific and quantitative control of a a genetic circuit output by means of a computationally designed, ligand dependent transcription factor and tissue specific promoter

2.1 Introduction

Designing synthetic genetic circuits to function within plant platforms can be extremely challenging. Plants evolved many layers of interconnected regulation required to perform necessary biological emergent behaviors. These behaviors include cell cycle progression, developmental programs, *e.g.*, progression through gametophyte or sporophyte generations, and cell differentiation to form various tissues and organs [1][2][3]. Underlying these emergent behaviors are molecular networks comprised of a plethora of biomolecules able to sense and adjust toward environmental change [4]. The sessile nature of plants has been a significant evolutionary pressure toward shaping traits to quickly and effectively respond to environmental conditions through plasticity of regulatory networks [5]. Therefore, regulatory networks evolved to be highly interconnected and complex within plants. Additionally, eukaryotes evolved genetic regulatory mechanisms beyond bacterial chassis such as alternative splicing, post translation modifications, and polyadenylation [6][7][8]. These mechanisms are tightly controlled, dynamic, and can change depending on the state of the cell at a given time. Ultimately, multicellularity arises from a variety of genes, the products of genes, and the interactions between them. The layers of regulation required for multicellularity can introduce issues and unpredictability when designing

synthetic genetic circuits [9]. However, multicellularity permits division of labor between cells, tissues, and organs and specificity of function(s) within them. For this reason, multicellularity can be wielded as a powerful trait toward programming different cell lines within engineered plants for downstream application.

The potential of engineered plants is vast despite the challenges. Plants could be engineered to produce biofuel, as biosensors integrated into agriculture crops, for commercial chemical and medicinal production, and as molecular tools to elucidate fundamental biological mechanisms and regulation [10][11][12]. Furthermore, human kind has relied on plant life for food, clothes, and shelter since the dawn of human civilization. Pertinent crops could be optimized to be more nutritive, to produce higher yields, to have greater photosynthetic efficiency, and to have reduced environmental impact [13][14][15][16].

A means to meet challenges associated with multicellularity is to impart specific and quantitative control of synthetic genetic circuity, *i.e.,* permit tunable output with more predictable behavior. This can be encoded with inducible and cell specific genetic circuits. If modularity is considered in genetic circuit design components could be easily modified or swapped, and a variety of inputs could be used to control the desired output within a given cell type. Furthermore, genetic circuits could be designed in such a manner as to permit interoperability to create whole systems [17].

Spatial control of gene expression can be genetically encoded into synthetic circuits with specific promoters (regulatory sequences of DNA upstream of coding regions). Tissue or cell specificity is dependent on the promoter driving expression of the desired gene. For example, Wolter et al., designed a heterologous Cas9 based genome editing tool targeting gene expression within egg-cells of *Arabidopsis thaliana* [18]. This system successfully increased the frequency of genomic edits by homologous recombination. This approach could be used to specifically express regulatory systems and genetic circuits in a desired cell type.

Genetic circuits composed of specific promoters and encoding regulatory proteins can temporally and quantitatively control gene expression, *i.e.,* output is dependent on external input. Specifically, some of these genetic systems have inducible promoters and cognate activators, *e.g.,* transcription factors (Figure 2.1) [19][20]. In some cases, if a promoter is regulated by an activator, addition of the ligand and binding of the activator to it causes the activator to stabilize. Consequently, stabilization permits the activator to bind the promoter activating transcription [21][22]. The circuit is “ON” or “OFF” in the presence and absence of ligand, respectively. Temporal control is achieved through addition of ligand, *i.e.,* transcription is activated or induced when the ligand is added. Furthermore, quantitative control is achieved by adjusting the ligand concentration which regulates the level of transcription.

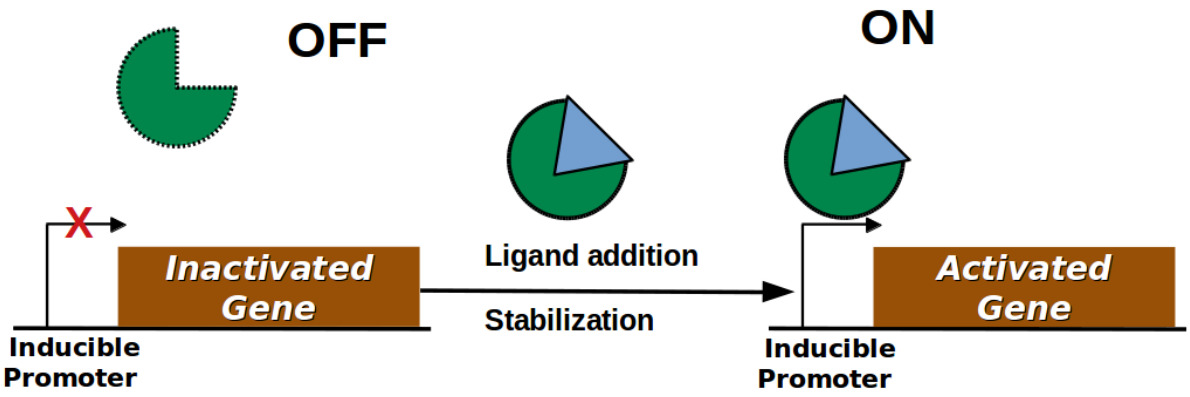


Figure 2.1. Promoters regulated by stability of activators. Transcription is inactive without ligand present (OFF). In the absence of ligand, the transcription factor is unstable and unable to bind the promoter. Transcription occurs with addition of ligand (ON). Ligand presence of absence regulates stabilization of transcription factor and in-turn regulates promoters.

An inducible system ought to be well designed to afford greater quantitative control of output. The activator should be specific for and highly sensitive to the inducer. Additionally, the system should be orthogonal to the host and have minimal or no basal output in the OFF state (termed leaky expression) [23][24]. Furthermore, it should have a large dynamic range, *i.e.,* a large disparity in the level of output in the ON and OFF states over a range of ligand concentrations.

A novel transcription factor developed through computational design with a degradation tag, and activation, DNA binding, and ligand binding domains has these characteristics [21]. The ligand binding domain was designed around the ligand, digoxin (DIG), consequently the domain has high specificity and sensitivity to DIG. High specificity results in orthogonality because the domain does not interact with other biomolecules in

heterologous hosts. High sensitivity affords quantitative control with DIG concentration [25]. Additionally, the ligand binding domain can be redesigned to bind other ligands, therefore this computational approach can be leveraged to create several biosensors [21][22][25]. Moreover, the transcription factor was computationally designed such that the ligand and DNA binding domains were conditionally stable [21]. In the absence of the ligand, the unstable transcription factor is degraded by the ubiquitin proteasome system, as a result the transcription factor is inducible and the system has minimal leaky expression. Furthermore, the transcription factor design was tuned to increase the dynamic range. Different activation domains were tested and compared to identify the domain creating the greatest dynamic range. To quantitatively control output within root epidermal cells, I developed a genetic circuit with the aforementioned genetic system (DIG transactivation system) and incorporated a root epidermal cell-specific promoter driving transcription of the activator to target output (Figure 2.2).

Cell-specific gene expression of the DIG system was achieved by encoding a root epidermal specific promoter (ABCG37) into the genetic circuit. In native context, the ABCG37 promoter drives transcription of an ATP-binding cassette transporter, PDR9, involved in auxin polar transport and root development [26][27]. In the synthetic circuit, the ABCG37 promoter is driving transcription of the computationally designed, ligand dependent transcription factor. This means the DIG transcription factor is transcribed

and translated within root epidermal cells, therefore throughout the whole plant body only root epidermal cells should respond to ligand and have output.

The DIG dependent transcription factor is composed of an N’-terminal Matα-degron tag fused a Gal-4 DNA binding domain (DBD), the DIG ligand binding domain (LBD), and a C’-terminal VP-16 activation domain [21]. As mentioned above, the DIG transcription factor is conditionally stable. Stability is achieved through binding DIG and subsequent dimerization of the DNA and ligand binding domains homodimers [21][25]. After becoming stable, the DBD can bind the inducible (UAS) promoter. The UAS promoter is composed of a Gal4 Upstream Activating Sequence (UAS) fused to a minimal (-46) Cauliflower mosaic virus 35S (CMV35S) promoter sequence. The UAS promoter is driving transcription of the output. In this study, the output is the reporter green fluorescent protein (GFP). GFP was used to qualify and quantify gene expression and dynamics of the gene circuit within roots. Spinning disk confocal microscope images were analyzed with an in-house MATLAB program to quantify GFP signal by means of pixel intensity.

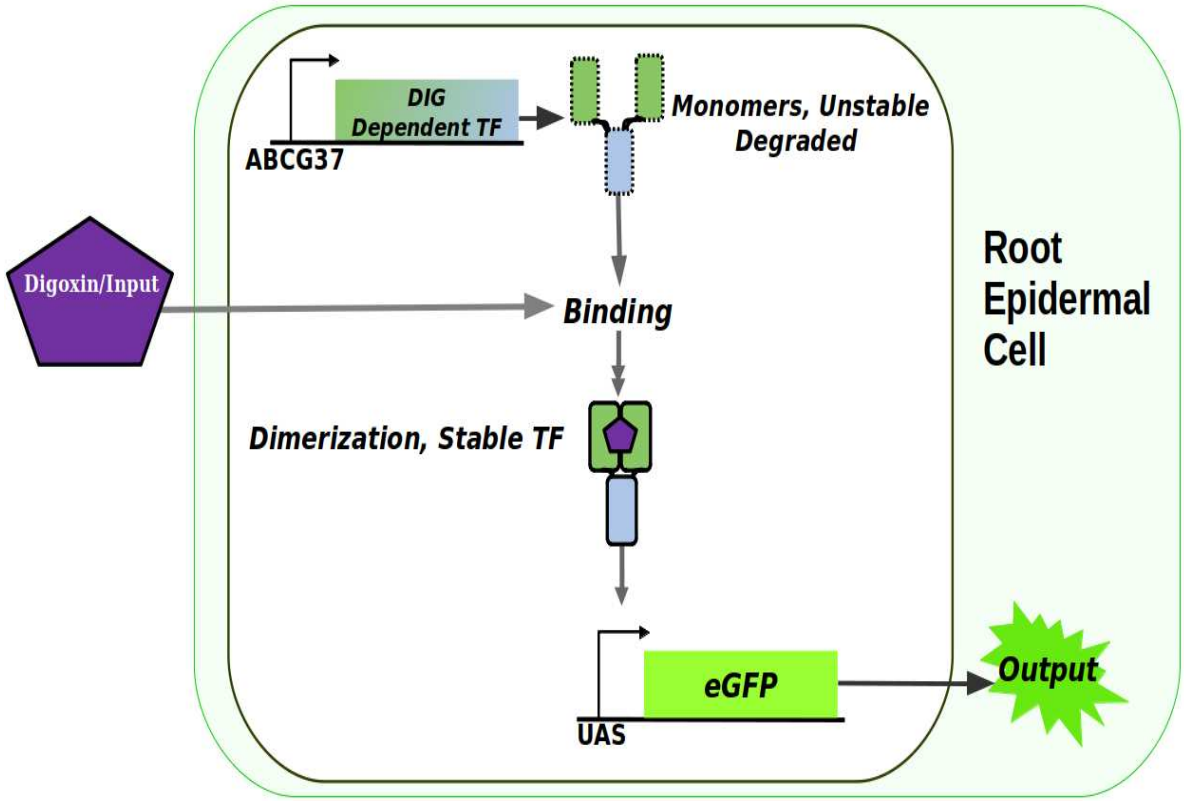


Figure 2.2. Inducible, root epidermal specific genetic circuit topology. Transcription of the computationally designed transcription factor (DIG dependent TF) is driven by the root epidermal cell-specific promoter (ABCG37). After addition of DIG, the DIG dependent TF can bind the inducible promoter (UAS). The UAS promoter drives transcription of output (GFP).

2.2 Methods

2.2.1 Assembly of the inducible, cell-specific genetic circuit

The UAS or inducible promoter (A.3) was amplified with Phusion high fidelity polymerase (NEB, Ipswitch, MA) using primers one and two (A.4) and pSEVA141-GFPmut 3.1 vector (Table 2.1) as template. EcoRI and AatII restriction enzyme sites were added to the 5’ and 3’ ends of the amplicon, respectively, using the aforementioned primers with the restriction enzyme sequences. PCR product was run on a 1% agarose electrophoresis gel to

verify correct size. The correct sized fragments were isolated from the agarose using a Zymoclean Gel DNA recovery kit (Zymo Research, Irvine, CA). The promoter fragment was directionally cloned into EcoRI and AatII (NEB, Ipswitch, MA) digested KJM340 (Table 2.2) using T4 ligase (NEB, Ipswitch, MA). The UAS promoter was cloned upstream of the GFP gene and downstream of the left T-DNA boarder in KJM340. The resulting plasmid was defined as SEO2 (Table 2.1). The 5’ end of the ABCG37 promoter and upstream region just beyond a PmeI restriction site were amplified with Phusion high fidelity polymerase using primers three and four (A.4) and KJM325 vector as template. The PCR product was screened and isolated from gel as above. By amplifying the 5’ end of the ABCG37 and upstream region, internal restriction sites, PmeI and XbaI (NEB, Ipswitch, MA), could be utilized to directionally clone the UAS and GFP sequences from into KJM325 to create SEO1 (Table 2.2). This required sub-cloning the upstream and 5’ ABCG37 promoter regions into SmaI and XbaI (NEB, Ipswitch, MA) digested SEO2, to generate SEO3 plasmid (Table 2.1). The DIG dependent TF was amplified with Phusion high fidelity polymerase with primers five and six (A.4) and the pSEVA-141-Matalpha-AtGal4-DIG10.3-VP16 vector as template. Fast digest BclI and DraI restriction enzyme sites were added to the 5’ and 3’ ends of the amplicon, respectively. The transcription factor fragment was directionally cloned into Eco53kI and BamHI (Thermo Scientific, Waltham, MA) digested KJM 325 vector, using T4 ligase, downstream of the ABCG37 promoter and upstream the NOS terminator to produce SEO4 plasmid. The UAS and GFP

sequences along with the 5’ end of the ABCG37 promoter were cut from the SEO3 vector using Eco53kI and XbaI (Thermo Scientific, Waltham, MA), and directionally cloned into PmeI and XbaI digested SEO4 using T4 ligase, downstream of the ABCG37 promoter and upstream of the NOS terminator to produce SEO1 (Figure 2.3). The final construct consisted of the inducible promoter (UAS promoter), driving the expression of the reporter (GFP), and the tissue specific promoter (ABCG37 promoter) driving the expression of the ligand dependent transcription factor (mat-alpha,Gal-4,Dig 10.3, VP16). The final construction was verified using NGS through Macrogen, USA (Rockville, MD). Samples with approximately 200 ng/µL of SEO1 plasmid and one of the primers from primers nine to 24 were sent in as reactions for sequencing.

Table 2.1: Templates for cloning

Construct Name	Description (5’→3’)	Forward Primer	Reverse Primer	Template sequence/PCR product	Added Restriction Sites (5’, 3’)
PSEVA 141-GFP mut3.1	UAS promoter:: GFP mutant:: TNOS::Txn block	1	2	UAS promoter	EcoRI, AatII
KJM325	ABCG37 promoter::Myb41 ::TNOS::Txn block	3	4	ABCG37 promoter 5’ fragment	PmeI, XbaI
PSEVA 141-Matalpha-AtGal4-DIG10.3-VP16	35S promoter::DIG dependent TF::OCS::Txn	5	6	DIG Dependent TF	BclI, DraI

Table 2.2: Vectors used and created to generate SEO1 plasmid

Purpose	Construct Name	Description (5'→3')	Restriction enzymes digested with	Removed fragments	Cloned fragments	Resulting construct from cloning
Sub clone of UAS promoter	KJM 340	Synthetic promoter:: GFP:: T-ADH::Txn	EcoRI, AatII	Synthetic promoter	UAS promoter	SEO2
Sub clone of ABCG37 5' fragment	SEO 2	UAS promoter:: GFP :: T-ADH::Txn block, ABCG37	SmaI, XbaI	N/A	ABCG37 promoter 5' fragment	SEO3
Source of UAS::GFP::T-ADH::Txn block and ABCG37' 5' fragment	SEO 3	UAS promoter:: GFP :: T-ADH::Txn block, ABCG37 promoter 5' fragment	Eco53kI and XbaI	UAS promoter:: GFP :: T-ADH::Txn block, ABCG37 promoter 5' fragment	N/A	
Sub clone of DIG dependent TF	KJM325	ABCG37 promoter::Myb41::TNOS::Txn block	BamHI, XbaI	Myb41	DIG dependent TF	SEO4
Final clone	SEO 4	ABCG37 promoter::DIG dependent TF::TNOS::Txn block	PmeI, XbaI	ABCG37 promoter 5' fragment	UAS::GFP::T-ADH::Txn block. ABCG37 promoter 5' fragment	SEO1

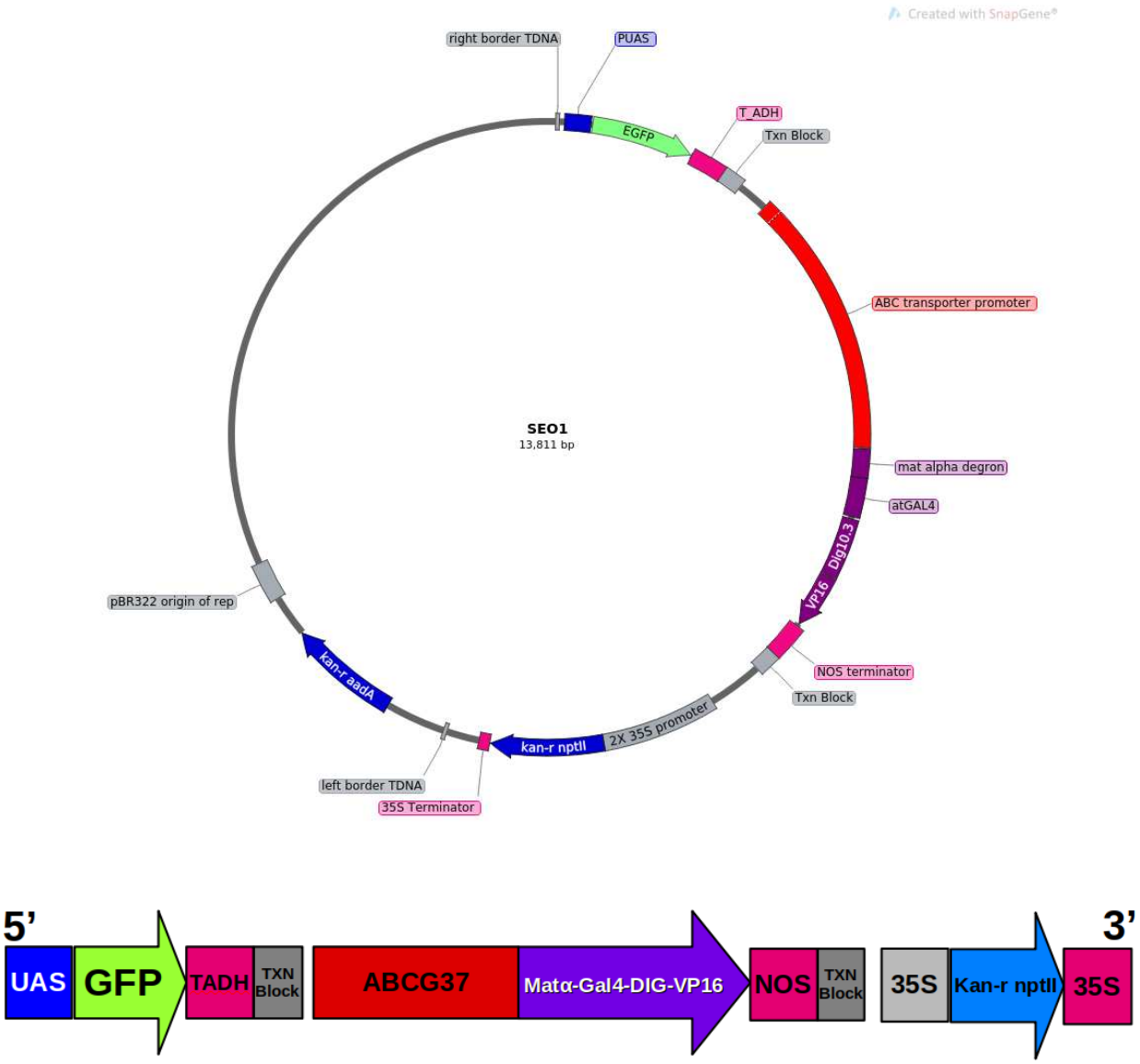


Figure 2.3. Vector map and T-DNA of inducible, root epidermal specific genetic circuit. A: Vector map, B: Linearized T-DNA The circuit is composed of two genes. From the left T-DNA boarder, or 5' end, is a UAS promoter driving transcription of enhanced GFP (eGFP), terminating in an alcohol hydrogenase terminator (T-ADH). In between this unit and the next is a transcription block to prevent read through and ensuring eGFP is transcribed and translated correctly. The root epidermal specific (ABCG37) promoter drives transcription of the digoxin (DIG) dependent transcription factor. This DNA unit terminates in a nopaline synthase (NOS) terminator. A transcription block is subsequent to the NOS terminator. Kanamycin markers were used for selection. Transcription of the kanamycin marker is driven by a 35S CaMV promoter and terminated by in an 35S terminator. Vector map figure was generated using SnapGene software.

2.2.2 Identifying stably transformed SEO1 Arabidopsis thaliana

After performing restriction cloning, competent Endura *E. coli* (NEB, Ipswitch, MA) strains were transformed using electroporation. Approximately 100 µL of *E.coli* was defrosted and 2 µL of SEO1 plasmid was added to the *E.coli.* The *E. coli* and plasmid mixture were placed in a cold cuvette (VWR, Radnor, PA). An ECM630-BTX (1250 v, 200 Ω,)was used to electroporate SEO1 plasmid into *E. coli*. The *E.coli* were suspended in 1 mL of fresh Luria-Bertani broth (LB) liquid media and recovered in a floor shaker set at 37°C and 200 rpm for approximately one hour. Approximately 200 µL of the recovered *E.coli* were grown overnight on LB media bacterial plates containing kanamycin (100 µL/mL). After growth, individual, isolated colonies resistant to kanamycin were further analyzed using GoTaq Green Master Mix (Promega, Madison, WI) with colony PCR. Colonies were screened for the UAS promoter using primers one and two, and pSEVA141-GFPmut 3.1 plasmid (Table 2.1) as a positive control. PCR product was run on a 1% electrophoresis gel.

Six *E. coli* colonies positive for the UAS promoter were grown overnight at 37°C and 200 rpm in approximately 2 mL of LB media containing kanamycin (100 mg/L). After growth, automated isolation of plasmid DNA was performed using a Qiaprep Spin Miniprep kit and Qiacube machine (Qiagen, Hilden, DE). Of the six plasmid samples two samples were utilized to transform competent GV3101 *Agrobacterium tumefaciens* via electroporation. Electroporation was performed as above. Agrobacteria were

grown for two days at 28ºC on bacterial plates containing LB media with kanamycin (100 mg/L), gentamycin (20 mg/L), and rifampicin (34 mg/L) selection. Two colonies were separately grown overnight in LB liquid media containing kanamycin (100 mg/L), gentamycin (20 mg/L), and rifampicin (34 mg/L). The cultures were placed in a floor shaker set at 28ºC and 200 rpm. One colony was selected to perform Agrobacterium-mediated transformation of *Arabidopsis thaliana* ecotype Columbia (wild-type plants) (A.2).

Selection of dipped plants, T0, was done by germinating seeds on Murashige and Skoog (MS) medium (A.1.2) with selection. Approximately 0.5 mL of seed from T0 plants was sterilized, and germinated on kanamycin (100 mg/mL) and cefotaxime (50 mg/mL) containing MS media plant plates. Seeds were vernalized for two days at 4 ˚C on plates. After which, plates were transferred to growth chambers. The growth chamber conditions were as follows: a photoperiod of 16 hours of light and 8 hours of darkness, at 22 ˚C, in 15% relative humidity. Plants were grown under these conditions for a minimum of 14 days. After 14 days of growth, kanamycin resistant plants were transferred to vertical plates, for root growth, containing MS media with kanamycin (100 mg/mL) and cefotaxime (50 mg/mL). Plants positive for antibiotic resistance and GFP signal were transferred to soil after screening and grown until seeds developed and matured. The seeds were harvested from positive T0 plants and were defined as the T1 generation.

T1 plants were initially screened for one T-DNA insertion event by kanamycin antibiotic resistance. Approximately 25 to 50 seeds of each T1

line were sterilized and germinated on kanamycin (100 mg/mL) containing MS media plant plates. After two days of vernalization at 4 ˚C, plates were transferred to growth chambers under the following conditions: a photoperiod of 16 hours of light and 8 hours of darkness, at 22 ˚C, in 15% relative humidity for a minimum of 10 days. Lines segregating a single locus according to Mendelian genetics, or 3:1 (resistant to sensitive), were screened for GFP signal. To screen for GFP signal, T1 plants were continually induced for 16 hours. Plants from lines positive for antibiotic resistance, segregating 3:1 (resistant to sensitive), and positive for GFP signal and inducible were transferred to soil and grown until seeds developed and matured. The seeds were harvested from the positive T1 plants and were defined as the T2 generation.

Table 2.3: Segregation and Chi-square data of SEO1 T1 lines

Line	Total Plants	Observed		Expected		X^2	Probability (p)
		Resistant	Susceptible	Resistant	Susceptible		
SEO1-2	48	35	13	36	12	0.1111	0.738
SEO1-21	20	18	2	15	5	2.4	0.12134
SEO1-45	50	38	12	37.5	12.5	0.026	0.87
SEO1-50	50	39	11	37.5	12.5	0.24	0.624
SEO1-51	50	34	16	37.5	12.5	1.3066	0.25299

T2 plants were initially screened for kanamycin antibiotic resistance. Approximately 50 seeds of each T2 line were sterilized and germinated on kanamycin (100 mg/mL) containing MS media plant plates. After two days of

vernalization at 4 ˚C, plates were transferred to growth chambers under the following conditions: a photoperiod of 16 hours of light and 8 hours of darkness, at 22 ˚C, in 15% relative humidity for a minimum of 10 days. Lines that did not segregate resistance to kanamycin selection were further analyzed.

2.2.3 Identifying inducible, root epidermal cell-specific GFP signal

T0, T1, and T2 populations were screened for GFP signal in root epidermal cells on a Lecia DM500 epifluorescence microscope with a 488 nm excitation and 525/30 nm emission filter set. To induce the genetic circuit, transgenic and wild-type plants were submerged in individual wells containing approximately 1 mL of liquid MS medium in the absence or presence of 100 µM DIG (Sigma Aldrich, St. Louis, MO) of a 24 well plate. The carrier for all DIG used was Dimethyl sulfide (DMSO) to ensure sufficient diffusion of DIG into cell (Thermofisher, Waltham, MA). After 16 (T1 and T2 plants) or 24 hours (T0 plants) of continuous induction, plants were washed in MS liquid media, roots were harvested, and subsequently mounted on microscope slides in 50% glycerol for observation under a microscope. Controls included wild-type plants grown in the same conditions simultaneously with other samples, and treated in the same manner as transgenic samples for each treatment. These controls are defined as wild-type controls. Transgenic plants treated with MS liquid media containing DMSO carrier at a final concentration of 1% were defined as transgenic controls. All mounted root samples were placed under excitation light for

200, 500, or 750 ms before imaging. Root samples from T0 and T1 lines with apparent, and robust GFP signal, and with a much greater GFP signal than controls were identified as positive. Plants with positively screened roots were transferred to soil, and grown until seeds developed and matured. The seeds were harvested from these transgenic lines and were defined as the T2 generation. A total of three T2 lines stably transformed with the inducible, root epidermal cell specific genetic circuit were analyzed for the first circuit.

A Keyence BZ-X700 epifluorescence microscope was used to observe GFP signal along the entirety of the root. T2 transgenic and wild-type plants seed were sterilized and germinated on vertical MS plates. Plants were grown for a total of 7 days. Upon the seventh day, plants were transferred from vertical plates to 24 well plates containing approximately 1 mL of liquid MS media, in the absence (DMSO) or presence of 100 µM DIG. After 16 hours of continuous treatment, roots were harvested and mounted on a microscope slide in 50% glycerol. BZ Analyzer software (Keyence, Itasca, IL) was used to produce composite images of the entire root. Additionally, this software was utilized to overlay composite fluorescent and bright-field images. Furthermore, measurements of images were taken with this software.

2.2.4 Confocal microscopy approaches to quantify GFP signal

To remove autofluorescence and background from confocal images, two spectrally separate emission filters and one excitation laser were used. Roots were exposed to 500 ms of 488 nm laser light to excite GFP. Emission filters with a 525 nm emission and 30 nm bandpass, and a 617 nm emission and

73 nm bandpass were employed to capture GFP signal and background, respectively (Figure 2.4).

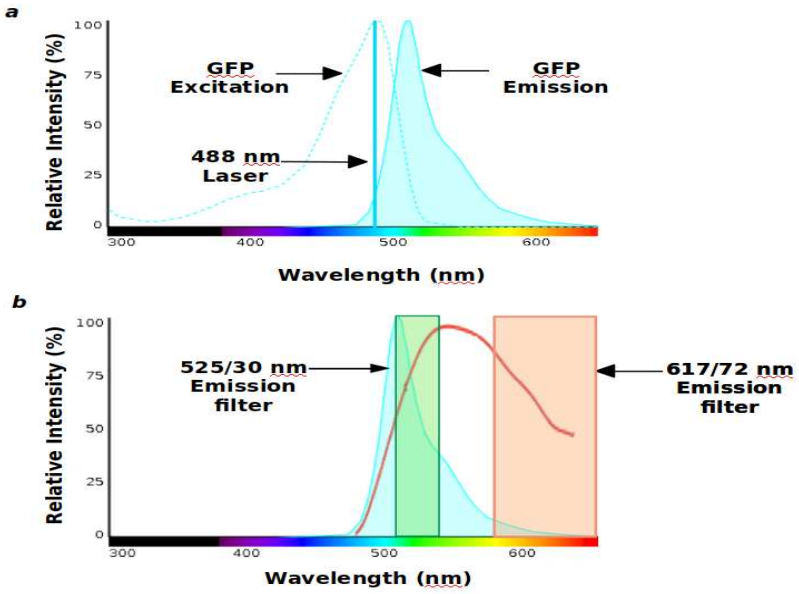


Figure 2.4. Excitation and emission spectra used in dual-wavelength confocal microscopy approach. a. 488 nm excitation laser is depicted with solid vertical blue line. GFP excitation and emission spectra are represented by the dashed and light blue shaded curves, respectively. **b.** Solid red line represents plant autofluorescence [26]. 525/30 and 617/73 emission spectra are represented by green and orange shaded blocks, respectively. The 525/30 emission filter spectrum represents capture of GFP signal and autofluorescence signal, *i.e.,* total signal. The 617/73 emission filter spectrum represents capture of autofluorescence. Figure was generated using Spectra-Viewer software (ThermoFisher, Waltham, MA).

A minimum of three points along each primary root within the maturation zone were taken with the confocal microscope for all quantitative experimentation, excluding images taken for development regulation experiments (Figure 2.3). Specifically, the zone of division and elongation are present once within the Arabidopsis primary root, *i.e.,* lateral roots were treated as a separate developmental class or zone. At each of the three

points, the root was optically sectioned over 80 microns, *i.e.,* a z-stack was generated for each point interrogated along the root (Figure 2.4).

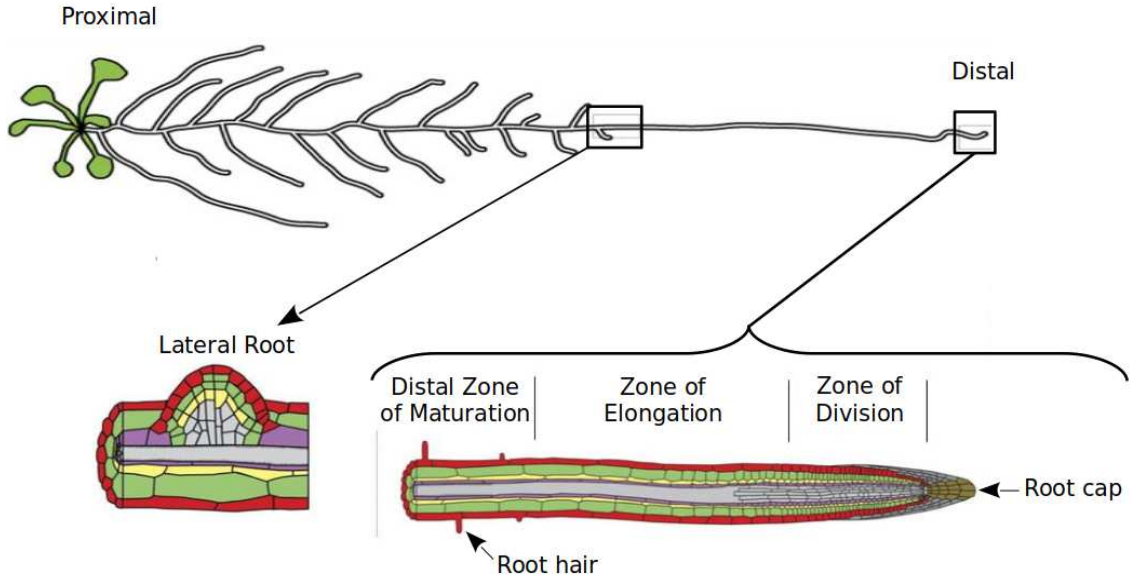


Figure 2.5. Developmental zones of the Arabidopsis root. The distal portion of the root begins with the root cap. Moving toward the proximal portion of the root, the zones are as follows: the zone of division, also known as the meristematic zone, the zone of elongation, and the zone of maturation, also known as the zone of differentiation. The proximal zone of maturation can develop lateral roots.

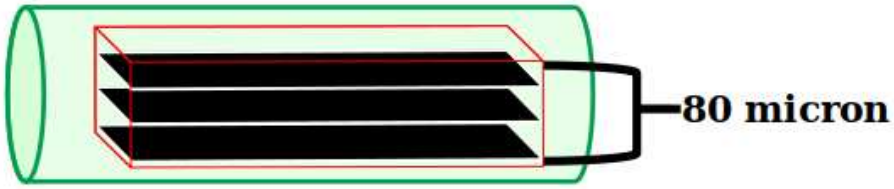


Figure 2.6. Optical sectioning of roots. The green cylinder represents a root. Each black rectangle represents an optical plane of the root within the zone of maturation. Two images were taken at each plane including one image using the 525 nm emission filter with a 30 nm bandpass (525/30) and the other using the 617 nm emission filter with 73 nm bandpass (6178/73). Each point of the root imaged was optically sectioned over 80 microns with five micron increments. All z-stacks were composed of 17 images, therefore a set was composed of 34 images.

2.2.5 MATLAB program to quantify level of GFP signal

Confocal images of wild-type roots were read into the program as a set of matrices, wherein each element represents a pixel at a given position. These matrices are run through a Pearson correlation, and linear regression, wherein the 617/73 (autofluorescence) emission filter images were the x variable and the 525/30 (GFP) emission filter images were the y variable. The Pearson correlation was run to ensure the autofluorescence within each channel was highly similar. Since the emission spectra of the filters are not identical, it is impossible to have identical autofluorescence signal captured in the images using these filters. However, if the images are highly correlated, above an (r) of 0.80, the similarity in autofluorescence is sufficient to predict autofluorescence in an image from a different image [28].

Assumptions must be met for linear regression including that the distribution of residuals must be normal. The assumption of residual normality is not required if the sample size is sufficiently large. This is because as sample size increases the distribution of residuals approximates a normal distribution [29]. The linear regression utilizes each pixel between the images at the same position therefore each pixel within the image represents a sample. Within one image there are approximately 1.56×10^5 pixels or samples, *i.e.,* the product of the dimensions (336 by 464 pixels) of each image. Furthermore, the sample size was expanded beyond the number of pixels within each image because multiple wild-type controls and technical

replicates of these controls were used for each linear regression. Moreover, each area imaged along the root was optically sectioned therefore a series of images was generated for each technical replicate (Figure 2.4). All of these approaches increased the sample size to create a sufficiently large sample size, well over the standard 30 samples to apply the Central Limit Theorem for linear regression [29].

The output equation from the linear regression was used to transform autofluorescence images. That is, the linear equation was applied to each pixel within the autofluorescence image. The linear coefficient physically represents the ratio of autofluorescence in each of the images. That is, the relationship defined through the linear coefficient is the amount of autofluorescence in each channel, wherein the dominator of the ratio is the autofluorescence within the autofluorescence 617 nm emission filter image and the numerator the autofluorescence of the 525 nm emission filter. The intercept represents the internal background arising from the confocal microscope, separate from the signal arising from autofluorescence of the root sample. The linearly transformed, autofluorescence images, represented by matrices, were directly subtracted from the corresponding total signal image pixel by pixel for each point interrogated on the root and for each optical section, *i.e.,* a pixel at a given position (x,y) within the autofluorescence image was subtracted from the pixel at that same position (x,y) within the total signal image. The MATLAB program subtracts pixel by pixel, by subtracting the matrices of the linearly transformed and total signal

images. This is performed, for each image set, over the entirety of the z-stack, therefore positional information within the image in the x, y, and z dimensions is maintained. The resulting image from subtracting these images contained GFP signal and excluded background signal.

Once background was removed, images were normalized from a scale of zero to one, for analysis in MATLAB, and logarithmically transformed. Images were transformed to increase contrast for more effective segmentation. Segmentation was performed by minimizing the weighted within-class variance to reduce a gray image to a binary image [30]. A bimodal histogram of the GFP image is produced (Figure 2.5). The valley between the two peaks is the threshold. Pixels under the threshold are zero and pixels above the threshold are one. Those pixels defined as zero are background. Pixels defined as one are the region of interest. The binary image overlays the original image as a mask. Pixels defined as zero in the binary image are also defined as zero in the original and are excluded from quantification. Pixels defined as one in the binary image are quantified, *i.e.,* the root sample with GFP signal within the original image is not occluded by overlaying the binary image. Subsequent to segmentation, the mean is calculated by taking the total sum of pixel intensities within the region of interest and dividing the sum by the number of pixels within the region for each mean. All images over all three z-stacks of a sample are taken to calculate the mean of the sample. These means are then used to calculate the grand mean of the treatment group. The pipeline is in Figure 2.6.

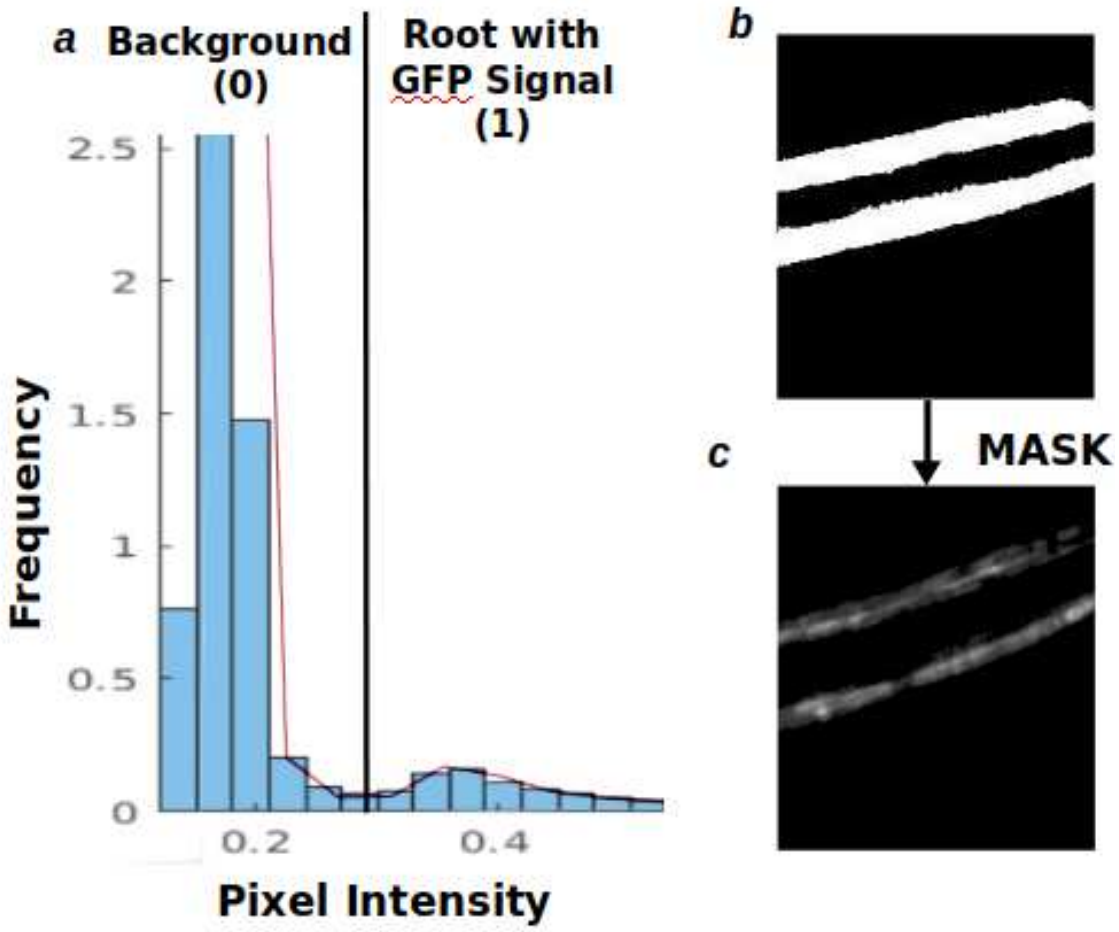


Figure 2.7. Histogram used to produce binary and masked images. a A histogram is generated for segmentation. The left peak represents the background and the right areas in the image with GFP signal within the root. The valley between the peaks represents the threshold. **b.** A binary image of a root. Pixels numbered zero, or those below the threshold, are black. Pixels numbered one, or those above the threshold, are white. **c.** The resulting masked image is the image without autofluorescence and background with the binary image acting as the mask.

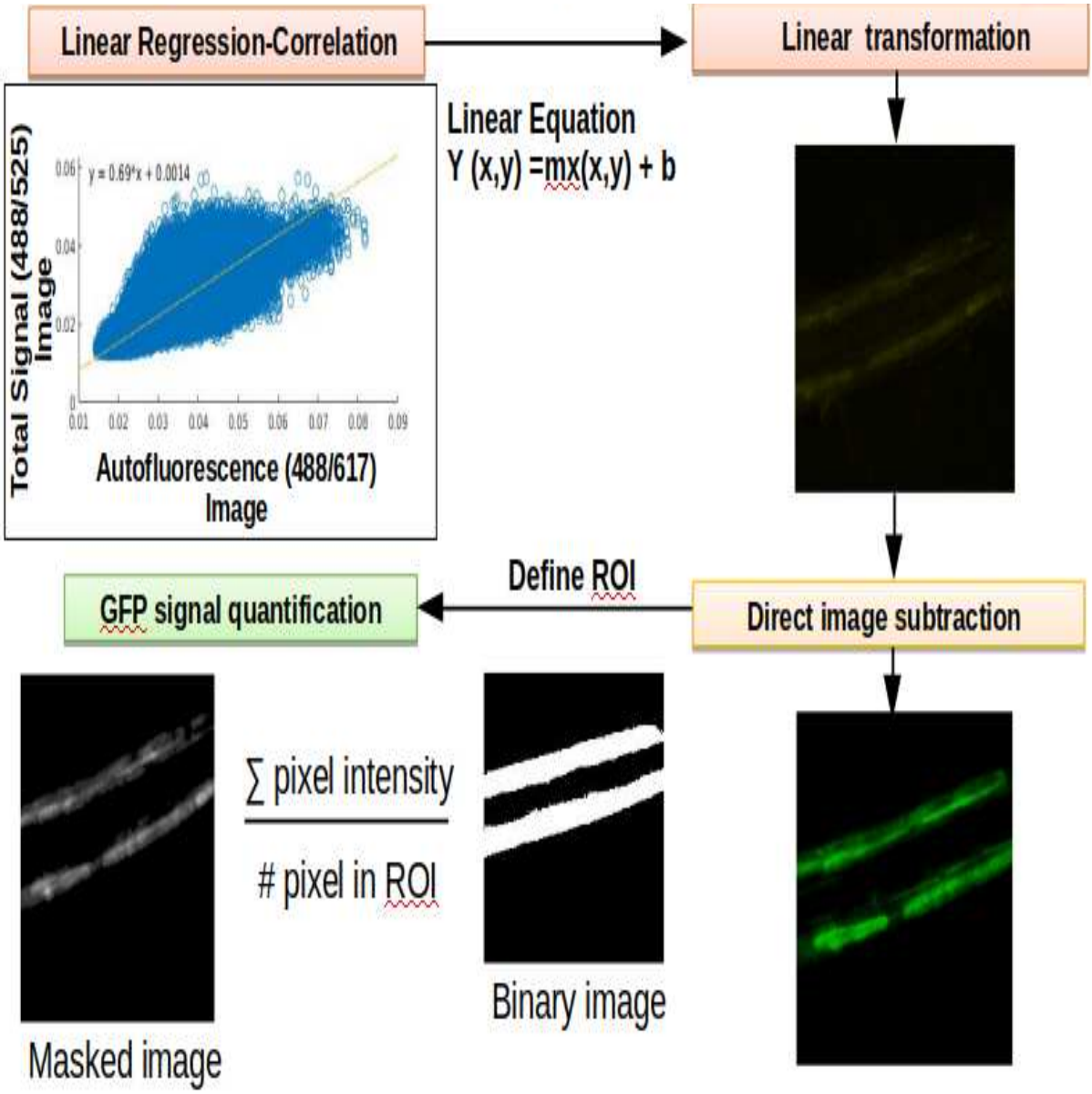


Figure 2.8. Pipeline of MATLAB program used to remove autofluorescence and background, and quantify GFP signal within an image. The program runs a linear regression between 617/73 confocal images (autofluorescence) and 525/30 (GFP signal) wild-type images. The linear equation from the regression is applied to the 617/73. The linearly transformed image is directly subtracted from the 525/30 image pixel by pixel to remove autofluorescence from the image. A region of interest is defined using Otsu's method. The mean pixel intensity is calculated throughout the region of interest.

2.2.6 Confocal microscopy and experimental methodology

Seed of T2 homozygous lines and *Arabidopsis thaliana* Columbia ecotype plants were sterilized and plated on vertical MS media plates containing kanamycin (50 µl/ml) antibiotic. After 10 to 14 days of growth under a light cycle of 16 hours of light and 8 hours of darkness, T2 and wild-type plants were transferred to 24 well plates containing liquid MS media and DMSO in the absence or presence of various DIG concentrations including 0.01, 0.1, 1, 10, 100, 200, and 500 µM. After 16 hours of continuous induction, plants were washed in liquid MS media, and roots were harvested. Harvested roots were mounted on glass slides in 50% glycerol and imaged using a Olympus IX83 Inverted Spinning Disk Confocal Microscope.

To identify if and when the genetic circuit was activated and in the “ON” state, and the point of maximal signal, T2 homozygous and wild-type control seeds were sterilized, plated, and grown as above. T2 plants and Arabidopsis controls were transferred to 24 well plates containing liquid MS media and DMSO in the absence or presence of 100µM DIG. Plants were removed from induction media after four and six hours, and 16, 24, and 40 hours of continuous induction, and quickly washed twice in fresh liquid MS media in the absence of induction media or DMSO carrier. These time points were interrogated to observe time of activation and maximal response, respectively. Preliminary screening was performed using a Lecia DM500 epifluorescence microscope with a 488 nm excitation and 525/30 nm emission filter set. From preliminary results, it was likely GFP signal was up

and beyond autofluorescence after six hours, and maximal after 16 hours of continuous induction. To quantify GFP signal, the experiment above was repeated with images of roots taken with an Olympus IX83 Inverted Spinning Disk Confocal Microscope with a 488 nm excitation laser using 525/30 nm and 617/73 nm emission filters. All samples were exposed to the 488 nm laser for 500 ms. Z stacks were taken at each interrogated point over 80 microns.

To determine the duration of the signal after removing induction media, T2 homozygous and wild-type seeds were sterilized, plated, and grown as above. T2 plants were transferred to 24 well plates containing liquid MS media in the absence or presence of 100 µM DIG. After 16 hours of continuous induction, all samples were washed by moving each plant to a clean well filled with fresh liquid MS media. The well-plates containing the plants were placed on a shaker for approximately one hour, after which plants were blotted dry. A total of two washes were performed. Subsequent to the two washes, plants were moved to vertical MS media plates. For preliminary work, plants were screened five, 18, 24, 72, 96, and 120 hours after being removed from induction media. Roots were harvested, mounted on glass slides in 50% glycerol. Images of the mounted roots were taken on a Lecia DM500 epifluorescence microscope with a 488 nm excitation and 525/30 nm emission filter set. From preliminary results, it was likely expression levels of controls and treatment groups were equivalent after 4 to 5 days post removal from induction media. To quantify GFP signal, the

experiment above was repeated. However, images of roots were taken on an Olympus IX83 Inverted Spinning Disk Confocal Microscope with a 488nm excitation laser using 525/30 nm and 617/73 nm emission filters. All samples were exposed to the 488 nm laser for 500 ms. Z stacks were taken at each interrogated point over 80 microns.

Preliminary analysis of epifluorescence images suggested GFP signal varied by developmental zone of the root. To identify if this was occurring, T2 homozygous and wild-type seeds were sterilized, plated, and grown as above. T2 plants were transferred to 24 well plates containing liquid MS media and DMSO in the absence or presence of 100 µM DIG. After 16 hours of continuous induction, all samples were washed with fresh MS media without inducer or DMSO present. Roots were harvested, mounted in 50% glycerol, and imaged with an Olympus IX83 Inverted Spinning Disk Confocal Microscope using 525/30 nm and 617/73 nm emission filters. All samples were exposed to a 488 nm laser for 500 ms to excite GFP. Z stacks were taken at each interrogated point over 80 microns. Interrogated areas of the root included distal and proximal portions within the zone of maturation, the zone of elongation, the zone of division, and lateral roots. Within each developmental zone, only one z-stack was taken for each root sample, unlike other experiments which interrogated three points within the zone of maturation along the length of the root. A total of six samples of each treatment were utilized for each independent transgenic line.

2.3 Results

2.3.1 Correlation and linear regression of wild-type root Images

For each experiment performed on the spinning disk confocal microscope wild-type roots were read into a custom MATLAB program. The program begins by running a Pearson correlation and linear regression wherein the autofluorescence (488/617) images are the x variable and the GFP (488/525) images are the y variable. The correlation between autofluorescence and total signal images was high, consistently near or above 0.86 (Figure 2.7).

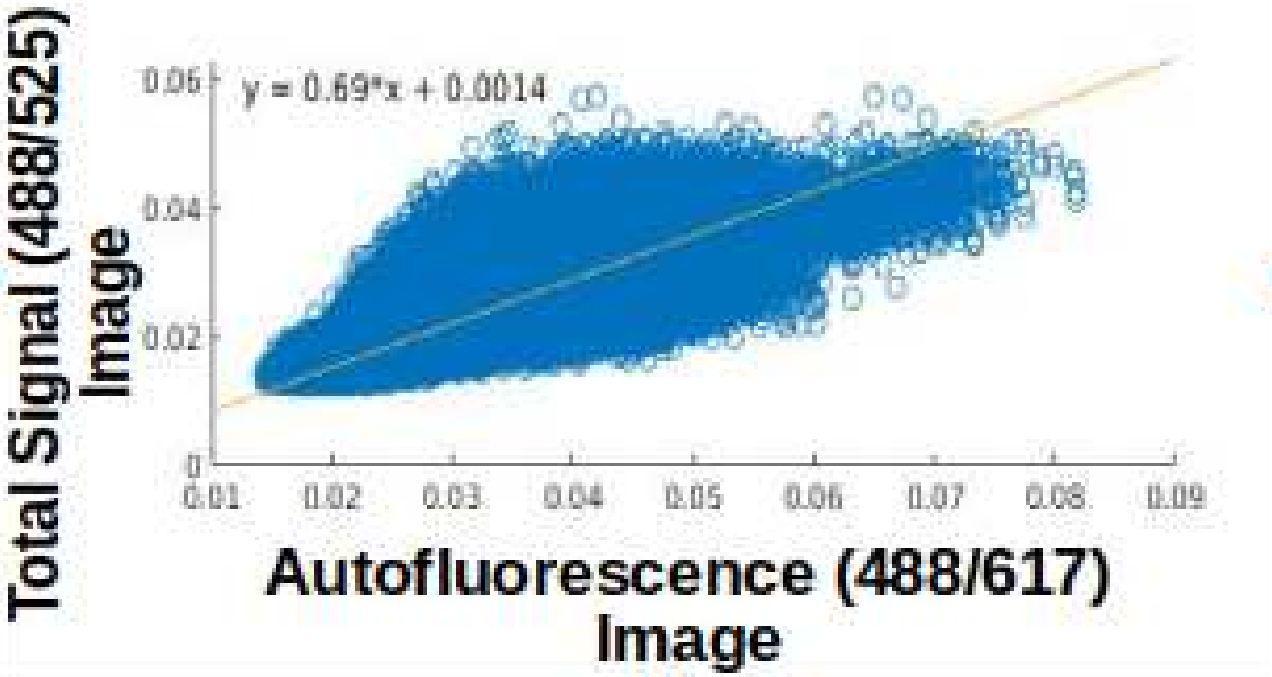


Figure 2.9. Representative Pearson correlation and linear regression of wild-type root autofluorescence and total signal images. x-axis: autofluorescence image pixel intensities; y-axis: Total signal image pixel intensities. The exact Pearson correlation *r* value for this model is 0.868. The linear equation is represented by the yellow line. The linear equation (y= 0.69x+0.0014) for the linear regression run on pixel intensities of the image is displayed in the upper left corner of the graph.

2.3.2 Root epidermal tissue-specific GFP signal

GFP signal was evident within root epidermal cells. Specifically, the outermost layer of the root was extremely robust in GFP signal (Figure 2.8). Furthermore, GFP expression was exclusively observed within the root epidermis layer, and extensions of the epidermis (root hairs) (Figure 2.8). Moreover, GFP signal was evident throughout the entirety of root within the root epidermal tissue layer (Figure 2.9). However, variation in GFP signal intensity was conspicuous. Specific developmental areas of the root epidermis exhibited strong GFP signal. These areas included the zone of division, the zone of maturation, and lateral roots (Figure 2.9, 2.10). Whereas other areas of the root epidermis exhibited weaker GFP signal, such as the zone of elongation (Figure 2.9, 2.10). The patterns of GFP signal within and between developmental zones observed in quantitative microscopy images follow those observed in epifluorescence images (Figure 2.11). The lateral roots of each independent transgenic line had an approximately two fold increase in GFP signal over the second most GFP signal intense developmental zone. These zones were the distal zone of maturation in the03 SEO1-45 line, and the zone of division in SEO1-50 and 21. Furthermore, GFP signal within the zone of maturation, either distal (SEO1-50 and 21) or proximal (SEO1-45), had a two and four fold increase, respectively, in GFP signal over the zone elongation.

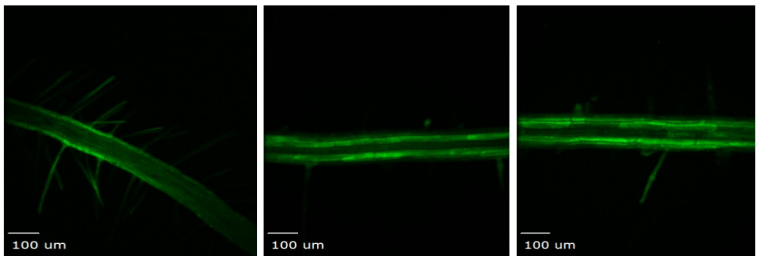


Figure 2.10. Confocal images of root epidermal specific expression. Representative transgenic roots treated with 1, 10, and 100 µM digoxin (left to right) for 16 hours. GFP signal can be seen in the outermost tissue layer (epidermal), and within the epidermal extensions (root hairs). All images where taken in the zone of maturation.

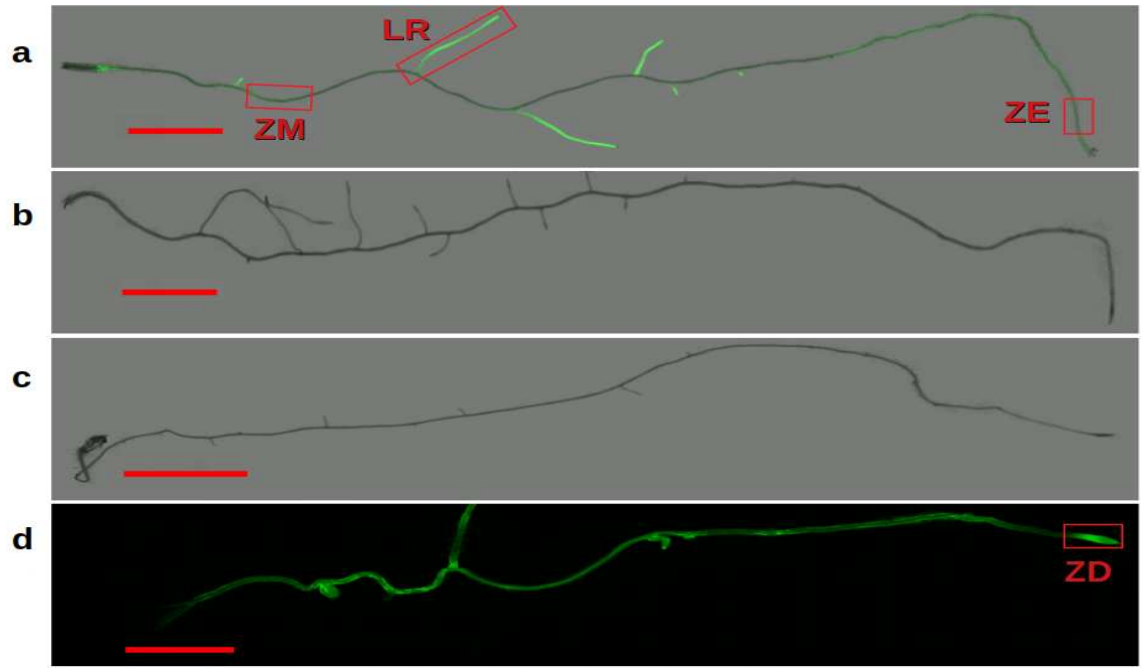


Figure 2.11. Whole-root composite images of entire roots. All images are representative. **(a)** Overlay of bright field and epifluorescence images of transgenic root sample after continuous induction for 16 hours with 100µM DIG medium. **(b)** Overlay of bright field and fluorescent images of transgenic control root. **(c)** Overlay of bright field and fluorescent images of wild-type control sample. **(d)** Fluorescence image of transgenic root under continuous induction for 16 hours 100µM DIG medium. Bars represent 5 mm. ZM: zone of maturation, LR: Lateral root, ZE: zone of elongation, ZD: zone of division.

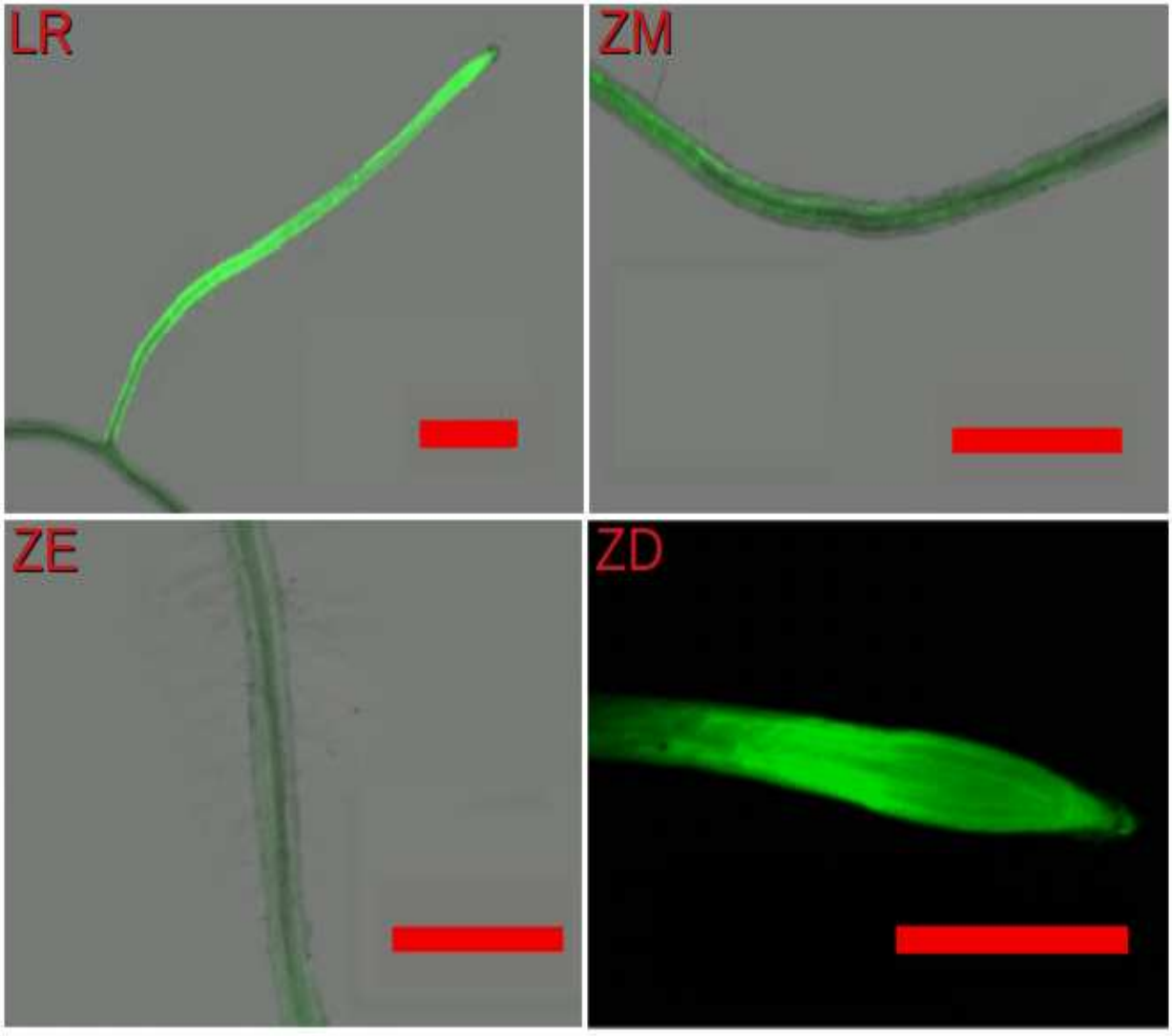


Figure 2.12. Close-up images of developmental . Close-up images taken from Figure 2.9. Upper left corner panel, lateral root (LR). Upper right panel, zone of maturation (ZM). Lower left panel, zone of elongation (ZE). Lower right panel, zone of division (ZD). The lateral root is extremely robust in GFP signal followed by the zones of division, maturation, and elongation. For zones of development refer to figure 2.5 and 2.11. Bars represent 1 mm.

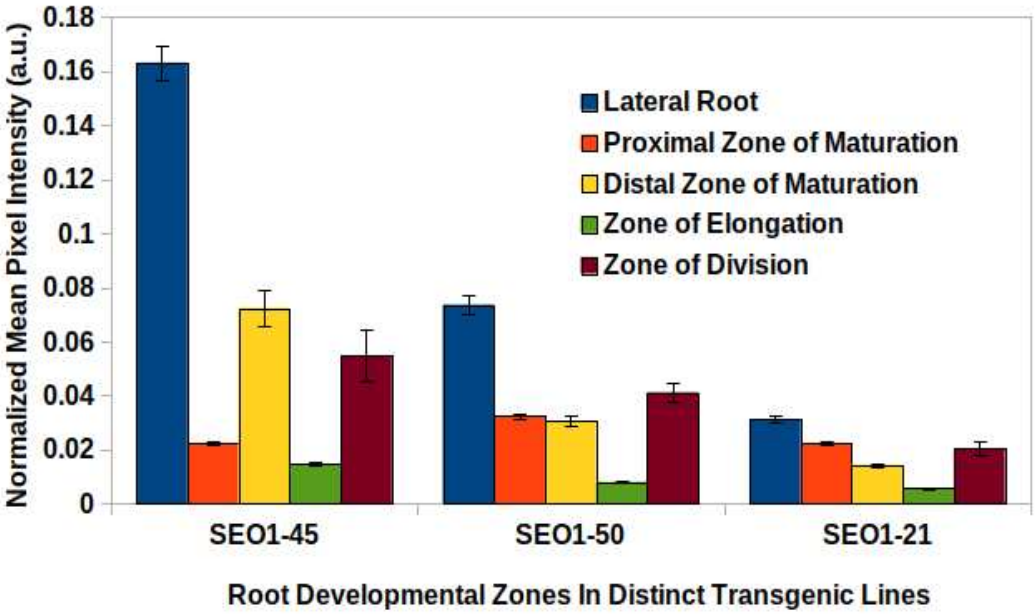


Figure 2.13. Quantified gene expression within each developmental zone of the root. Transgenic plants and wild-type controls were treated with 100µM DIG and continuously induced for 16 hours. Interrogated zones of the root included the zones of division, elongation, and maturation, and lateral roots. All images of treated transgenic samples have been normalized to transgenic controls to account for basal expression. Error bars represent +/- 2 SE. n=6 per treatment.

2.3.3 Induciblity: Activation and duration characterization

Transgenic controls, those not exposed to ligand and otherwise treated identically to treated transgenic samples, had no discernible GFP signal above autofluorescence. Transgenic controls appeared to have fluorescence signal consist with wild-type controls (Figure 2.9).Transgenic roots exposed to DIG had extremely evident fluorescence signal within root epidermal cells and extensions of the root epidermis (root hairs) in epifluorescence images. Furthermore, the basal expression of the lines was extremely low. SEO1-45 had the greatest level of basal expression, nearly 3% of that observed in the

induced sample (Figure 2.12). The other two lines did not have as high of basal expression. Nearly 2.5% for SEO1-21, and 1% for SEO1-21 of the induced sample.

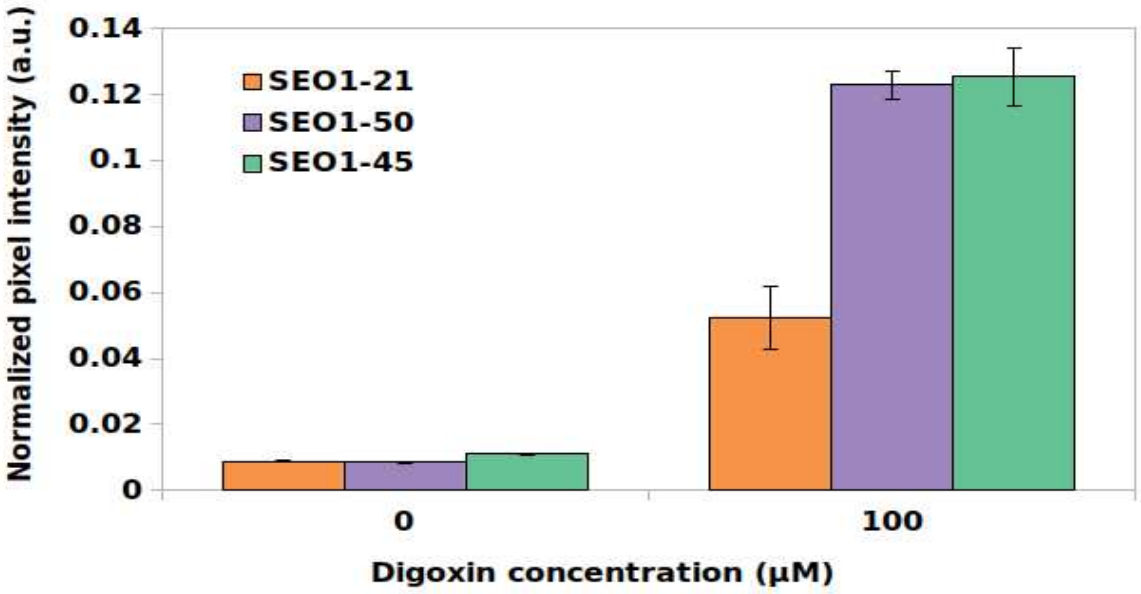


Figure 2.14. GFP signal under no and digoxin media. Transgenic controls are represented by the zero concentration. Samples were treated with treatment media continually for 16 hours. Error bars represent +/- 2 SE. n=6 per treatment, three technical replicates per sample.

After six hours of continuous induction fluorescence above transgenic controls was observed in treated samples (Figure 2.13). Furthermore, roots continuously induced for 16 hours had the greatest observed signal intensity in all independent transgenic lines over all observed time points. GFP signal was less at 24 hours than at the peak observed at 16 hours. Both SEO1-21 and 45 transgenic lines had just over a four fold increase in signal, and SEO1-50 had just under a seven fold increase. However, GFP signal was greater after 24 hours of continual induction than at six hours of continual induction

by nearly three fold in SEO1-50, four fold in SEO1-21, and 4.6 fold in SEO1-45. By 40 hours, GFP signal increased from the decline observed after 24 hours of continual induction. SEO1-21 GFP signal was nearly the same at 24 and 40 hours, whereas GFP signal was greater in SEO1-50 and lesser in SEO1-45 by 15.5% and 33%, respectively.

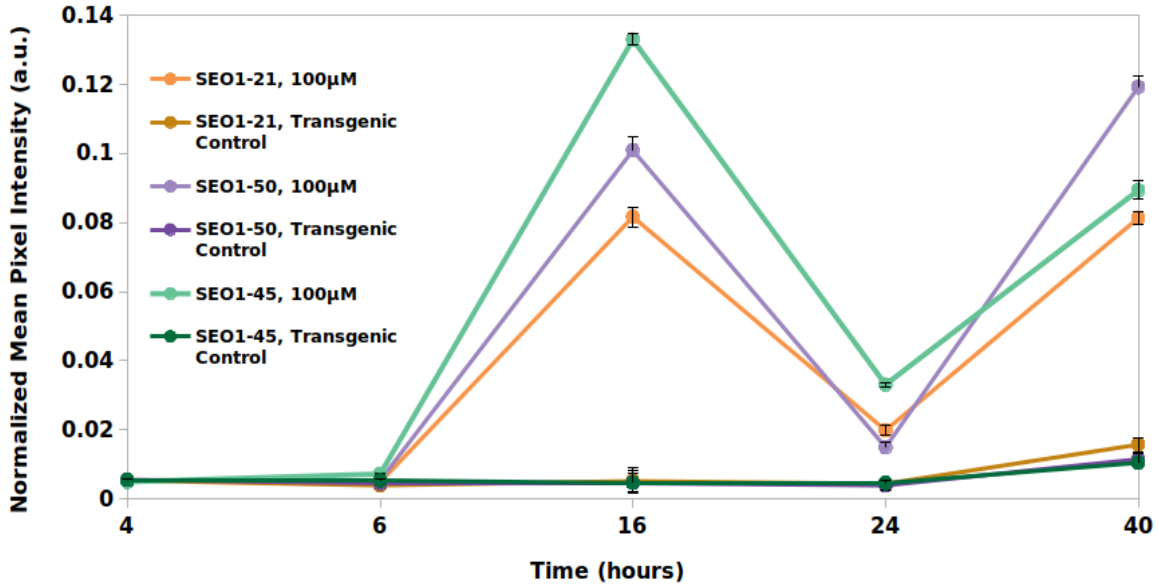


Figure 2.15. GFP signal over time of continually induced transgenic samples. Control and transgenic samples were continually induced for four, six, 16, 24, and 40 hours. At each time point, transgenic and wild-type controls were also sampled. Error bars represent +/- 2 SE n= 6 per line, three technical replicates per sample.

After removal of induction media GFP expression decreased overtime until no discernible GFP signal was observed. 24 hours after being removed from induction media GFP signal intensity decreased by nearly 56% in SEO1-21 samples, and nearly 59% in SEO1-45 and SEO1-50 samples compared to samples continually induced for 16 hours (Figure 2.14). Five days following removal from induction media, GFP signal did not significantly differ within

root epidermal cells between transgenic controls and treated transgenic samples in all lines (Figure 2.15). Interestingly, the decline of SEO1-21 GFP signal was slower than SEO1-45 and SEO1-50. In particular, the decrease between 24 and 72 hours, wherein the slope of the line of SEO1-21 samples is far less steep.

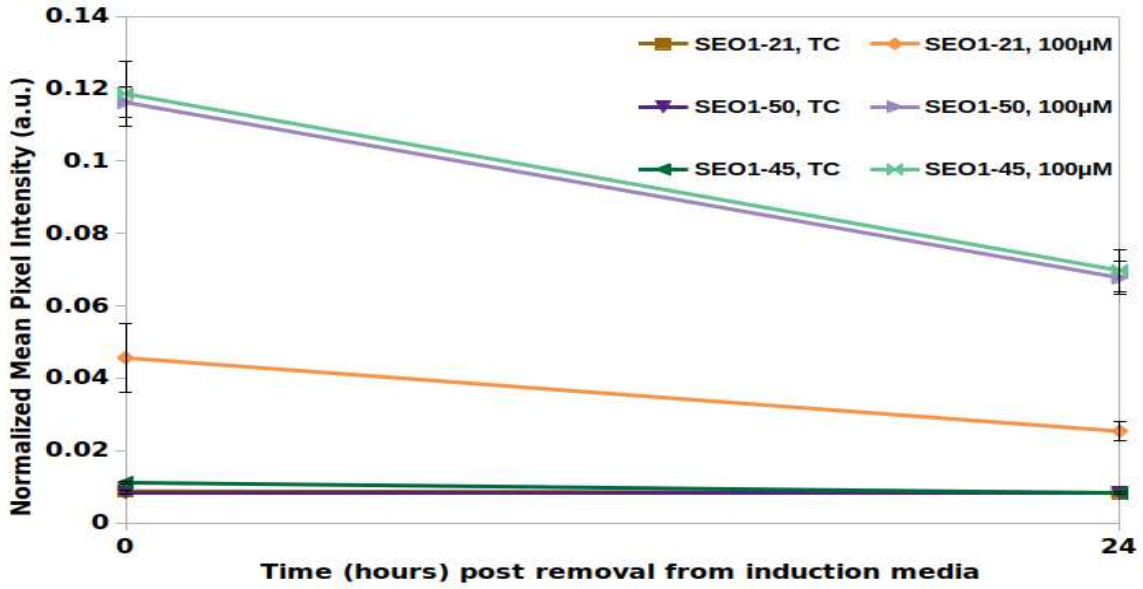


Figure 2.16. GFP signal in independent transgenic lines 24 hours post removal from induction media. Continually induced transgenic samples (16 hours) represent time zero, *i.e.,* these samples were removed from induction media and immediately screened. Samples were continually induced for 16 hours, washed, and screened 24 hours after removal from induction media are represented by the 24 hour time point. n= 5 per line, three technical replicates per sample. Error bars represent +/- 2 SE.

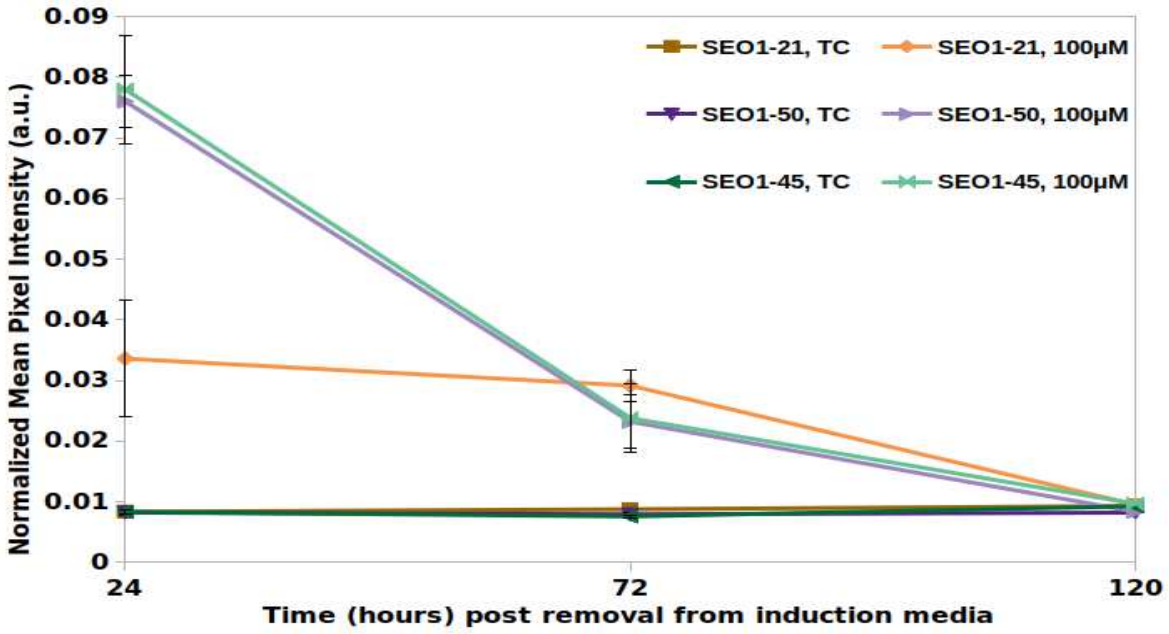


Figure 2.17. Duration of GFP signal after removal from induction media. GFP signal 24, 72, and 120 hours after removal from induction media. The GFP signal intensity of all transgenic lines 24 and 72 hours after removal from induction media significantly differed from transgenic controls. All treated samples did not significantly differ from transgenic controls after 120 hours post removal from induction media. n= 5 per line, three technical replicates per sample. Error bars represent +/- 2 SE.

2.3.4 Dose-curves and quantitative control

As the concentration of ligand increases, the level of GFP signal increased. A minimum ligand concentration of 100nM was required to observe significantly different GFP signal from transgenic controls (Figure 2.16). Additionally, ligand concentrations above 100μM produced nearly equivalent or less GFP signal (Figure 2.16).

Transgenic root samples under continuous induction at a 100μM digoxin concentration had a 34 to a 150 fold increase in GFP signal relative to

transgenic controls (Figure 2.17). Of the three lines tested, two had fold-changes of approximately 34 and 43. The lines with lowest and highest absolute GFP signal, SEO1-45 and SEO1-21, respectively, had smaller fold inductions relative to line with an intermediate GFP signal, SEO1-50. This line had an observed fold induction in signal intensity nearly 3.5 to 4.5 times over the other lines. The dynamic range of these lines mirrored the fold induction. For example, SEO1-50 samples had the greatest fold induction and the greatest dynamic range.

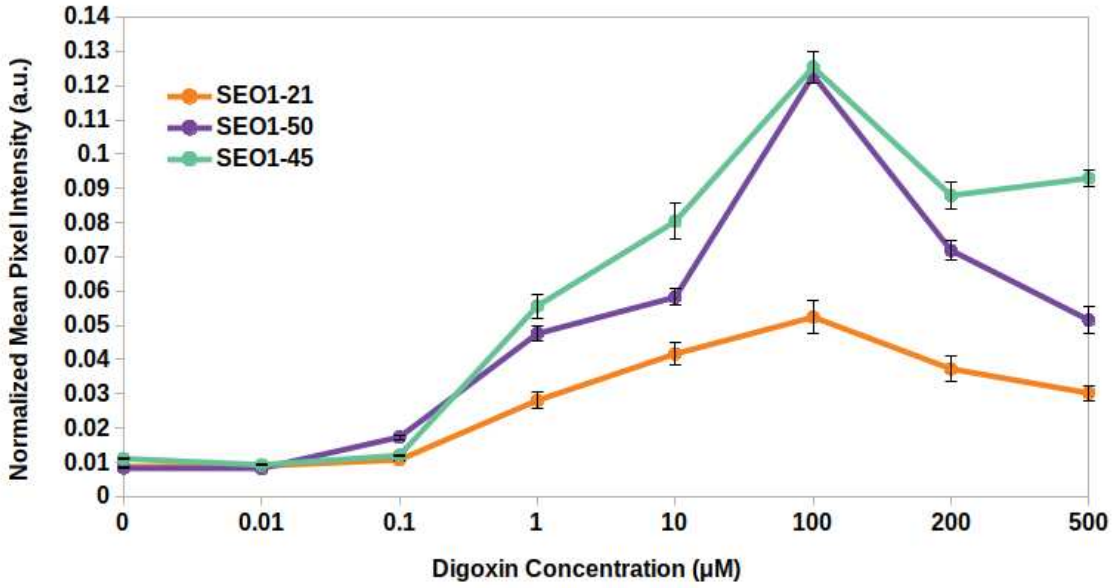


Figure 2.18. Dose-dependent response of independent transgenic lines. Three independent lines (SEO1-21, SEO1-45, SEO1-50).Treated samples were continuously induced for 16 hours over eight concentrations of digoxin including 0, 0.01, 0.1, 1, 10, 100, 200, and 500 µM. All samples have been normalized to wild-type control. Error bars represent +/- 2 SE. n=6 per treatment, three technical replicates per sample.

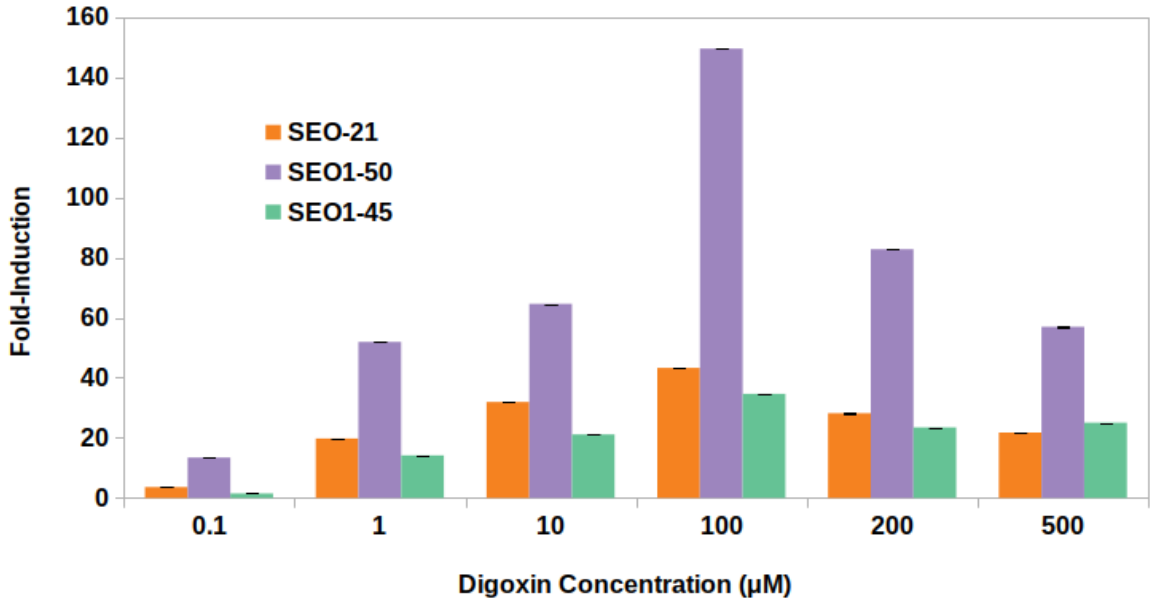


Figure 2.19. Fold-induction of individual transgenic lines over several ligand concentrations. Fold induction was calculated by normalizing to wild-type sample and calculating the ratio of GFP signal between induced transgenic samples and transgenic controls. Error bars represent 95% confidence interval. n=6 per treatment.

2.4 Discussion

An inducible, tissue-specific genetic circuit permits spatiotemporal and quantitative control of a desired output. Here, I have demonstrated an inducible, root-epidermal specific genetic circuit stably integrated *in plantae.* To quantify GFP signal, an experimental approach using one excitation laser with two emission filters was developed. Images output from the experimental approach were subsequently fed into a computational pipeline for high-throughput analysis (Figure 2.8).

A challenge in characterizing genetic parts and circuits by means of fluorescent reporters and microscopy is quantifying reporter signal in a high-throughput manner. This is particularly true when imaging living tissues

which emit autofluorescence that can partially occlude “true” fluorescence reporter signal. To meet this challenge, I developed an experimental approach with dual-wave length emission filters and incorporated a custom MATLAB program into the pipeline. The computational approach relied on linear regression between autofluorescence and GFP images necessitating high correlation between images of a set, which was observed across experiments. The minimum Pearson correlation (*r)* between sets of images was just above 0.86 ranging to nearly 0.94. This is well above the minimum correlation required to accurately model image sets [26][27]. In addition to being highly correlated linear equations varied little across experiments. The range of the slope was between 0.41 and just over 0.65, with a minimum intercept of 0.0016 and maximum of 0.0078. Having observed consistently high correlation between images and minimal variation between equations across experiments suggests linear regression accurately captures the autofluorescence and background signal relationship between the emission filters as to accurately remove these signals.

Relying on linear regression to quantify fluorescence reporter signal does introduce a limitation being the availability of spectrally separated emission filters able to capture both autofluorescence and reporter signal simultaneously. However, confocal microscopes are designed in such a manner to have a variety of excitation lasers and emission filters. Moreover, a variety of fluorescence reporters are available to chose from. This widens the field of possible and compatible excitation wavelengths and emission

filters to the reporter. Furthermore, autofluorescence excitation and emission spectra are extremely broad, therefore autofluorescence can be excited and emission signal captured utilizing most lasers and filters. With a wide range of available excitation lasers and emission filters that can fit desired experimental and computational approaches this limitation is readily overcome.

Transcription of the digoxin dependent transcription factor driven by a tissue-specific promoter imparted spatial control over output of the DIG system. The spatial patterns of GFP signal did align with characterization studies of the ABCG37 promoter [32]. Moreover, spatial patterns of GFP signal aligns with expression patterns of the native ABC transporter (ABCG37/PDR9) [33][34]. GFP signal occurred exclusively within root epidermal cells throughout the entirety of the root, and intensity of GFP signal had a pattern dependent on the developmental zones of the root. For instance, lateral roots had greater overall signal, whereas the elongation zone consistently had the lowest overall signal. within the epidermal tissue layer compared to other developmental zones of the root. In general, regulation and spatial expression patterns of a promoter driving transcription of the DIG dependent transcription factor extends to the output of the DIG system. While the promoter driving transcription of the DIG dependent transcription factor spatially regulates output, temporal regulation of it is controlled by the ligand dependent transcription factor.

The DIG dependent transcription factor behaves in an inducible manner with an “ON” state and “OFF” state in the presence and absence of ligand, respectively. Leaky expression in the OFF state is minimal. The output of the leakiest line approximated 3% of the maximal output. The other analyzed transgenic lines, with less leaky expression, had outputs approximating 1% and 2% of the maximal response. This indicates the DIG system is tight, with a low level of basal expression. Additionally, the sensitivity, or ligand concentration required to induce the DIG system and enter it into an ON state, is equal to the dexamethasone, a synthetic glucocorticoid, inducible system (pOp/LhGR) [35]. Both required a ligand concentration of 0.1μM to activate. However, the sensitivity of the widely studied β-estradiol inducible system (XVE/OlexA), 0.008μM, is far less than these systems [36]. Although, the fold-induction of the XVE/OlexA system is nearly 18 times less than the DIG system.

The difference between leaky expression and maximal output of DIG transgenic lines represents a reasonable dynamic range. A 150 fold-induction was observed in the transgenic line with the largest dynamic range. In comparison, the XVE/OlexA with a fold-induction of eight represents a far smaller dynamic range [36]. The pOp/LhGR and alcohol inducible (AlcR/AlcA) systems have far greater fold inductions relative to the DIG system, of 1000 and 2000, respectively [35][37]. The dynamic range of the DIG system is neither extremely low nor high compared to other systems, but the ligand concentration required for maximal response is greater. For example, ligand

concentrations of 10 and 0.08uM elicited maximal responses from the pOp/LhGR and XVE/OlexA systems, respectively. Although, these systems require a lower ligand concentration to reach maximum output, this lends to a smaller range of ligand concentrations eliciting response. The smaller range could limit output tunability. The fold-induction is extremely high and the range of responsive ligand concentrations small, therefore small adjustment to ligand concentration would produce a larger change in output from these systems relative to the DIG system. This should be taken into consideration if output needs to shift on a finer-scale. For instance, output from the genetic circuit may feed into another transcriptional pathway wherein the outcome of the downstream pathway is dependent on a gradient. In Arabidopsis cell fate and differentiation, developmental trajectories of tissues and organs, and defense mechanisms are often determined by gradients of phytohoromones and reactive oxygen species [38][39][40]. In-turn, gradients are regulated and established by regulatory networks and transcription factors composing those networks. To establish and regulate a gradient to ensure a desired outcome regulatory network outputs need to be adjusted accordingly.

Toxicity issues have been observed with the glucocorticoid and ethanol inducible systems. The rat receptors incorporated into glucocorticoid inducible systems as GR domains can negatively effect development [33]. Furthermore, exposing plants to ethanol, particularly in the context of the AlcR/AlcA system, has been shown to reduce viability [41]. These issues were

not observed within plants transformed with the DIG system. Toxicity may occur with high exogenous DIG concentrations. However, toxic effects were observed as reduced output, and no development nor viability consequences were observed. Additionally, more concentrated ligand can be avoided as the maximum output occurs below possible toxic ligand concentrations.

The time to required to activate the DIG system, for output to reach a maximal response, and the duration of signal after removal from induction media align with other inducible systems. Both the pOp/LhGR and DIG systems had observable output after six hours of continuous induction [6]. It should be noted, the pOp/LhGR system was within a tissue-specific context, specifically the quiescence center of the root meristem. However, a one hour activation time has been reported for the dexamethsone system not within a tissue-specific context [38]. Additionally, XVE/OlexA and AlcR/AlcA systems had observable output after one half and four hours, respectively [36][37]. Although these systems turn ON earlier, the DIG system reaches maximal response faster. The XVE/OlexA and pOp/LhGR reach maximal response after 24 hours of continuous induction, whereas the AlcR/AlcA system requires five days of continuous induction. Temporal differences between these systems may relate to molecular mechanisms underlying them. Among many possibilities are binding affinity of transcription factors to DNA regulatory sequences and ligands, interactions with transcriptional machinery, and degradation rates of system mRNA and ligand [39][40][41]. Additionally, the pOp/LhGR relies on spatial separation of the LhGR transcription factor from

the cognate DNA, therefore nuclear import may retard progression toward maxima. Whatever the case, the DIG system reaches the maxima faster.

The DIG output increases and decreases in what appears to be a circadian rhythm. This could relate to the ABCG37 promoter and its association with auxin transport. Regulation of auxin signaling and responses by the circadian clock, and other circadian behaviors regulated by auxin have been shown [46][47]. Temporal studies analyzing the DIG system with quantitatively and well-characterized promoters may elucidate if a circadian rhythm arises from innate characteristics of the system or parts of the system, or as a result of the promoter driving transcription of the DIG dependent transcription factor.

The genetic circuit I designed is modular, therefore can be easily manipulated in variety of ways in future studies as a molecular tool. For instance, as a molecular tool to functionally and quantitatively characterize promoters. Moreover, the correct topology and design of the genetic circuit with CRISPR/Cas9 gene editing tools can spatially and temporally control editing events. For instance, if the UAS promoter drives transcription of the Cas9 protein the editing tool would be inducible, temporal control, and possibly quantitatively controlled by the ligand concentration. Furthermore, CRISPR/Cas9 off-target editing events could be limited to a subset of cells or developmental stage by driving transcription of the DIG transcription factor by a tissue or cell specific promoter. This could limit negative effects caused by off-target effects throughout the whole plant body.

The DIG system could also be incorporated into metabolic engineering designs. Particularly if a biopolymer or metabolic pathway was required in a specific cell-type or tissue and needed to be temporally controlled. Economically important tomato crops, specifically the fruit, could be engineered and used as a biotechnological resource for medicinal bioactive products [19]. The DIG system could be used to activate metabolic pathways in the fruit to produce these medicinal and bioactive products at the desired time, particularly before ripening occurs and the metabolic demand is low.

The inducible, tissue-specific circuit is an additional genetic system of the Arabidopsis molecular toolkit. The system can provide an approach to expand the library of characterized genetic parts, be utilized in metabolic engineering, and impart spatiotemporal tunability of genetic expression.

REFERENCES

[1] Jones, A. R., Forero-Vargas, M., Withers, S.P., Smith, R.S., Traas, J., Dewitte, W., & Murray, J.A.H. “Cell-Size Dependent Progression of the Cell Cycle Creates Homeostasis and Flexibility of Plant Cell Size.” *Nature Communications*, vol. 8, no. 1, Apr. 2017, doi:10.1038/ncomms15060.

[2]Sakakibara, K., Ando, S., Yip, H.K., Tamada, Y., Hiwatashi, Y., Murata, T., Deguchi, H., Hasebe, M., & Bowman, J.L. “KNOX2 Genes Regulate the Haploid-to-Diploid Morphological Transition in Land Plants.” *Science*, vol. 339, no. 6123, 28 Mar. 2013, pp. 1067–1070., doi:10.1126/science.1230082.

[3] Drapek, C.E., Sparks, E.E., Benfey, P.N. “Uncovering Gene Regulatory Networks Controlling Plant Cell Differentiation.” *Trends in Genetics*, vol. 33, no. 8, Aug. 2017, pp. 529–539., doi:10.1016/j.tig.2017.05.002.

[4] Kollist, H., Zandalinas, S.I., Sengupta, S., Nuhkat, M., Kangasjärvi, J., & Mittler, Ron. “Rapid Responses to Abiotic Stress: Priming the Landscape for the Signal Transduction Network.” *Trends in Plant Science*, vol. 24, no. 1, 3 Nov. 2018, pp. 25–37., doi:10.1016/j.tplants.2018.10.003.

[5] Lachowiec, J., Queitsch, C., & Kliebenstein, D.J. “Molecular Mechanisms Governing Differential Robustness of Development and Environmental Responses in Plants.” *Annals of Botany*, vol. 117, no. 5, 14 Apr. 2015, pp. 795–809., doi:10.1093/aob/mcv151.

[6]James, A.B., Syed, N.H., Bordage, S., Marshall, J., Nimmo, G.A., Jenkins, G.I., Herzyk, P., Brown, J.W.S., & Nimmo, H.G. “Alternative Splicing Mediates Responses of the Arabidopsis Circadian Clock to Temperature Changes.” *Plant Cell*, vol. 24, no. 3, 2012, pp. 961–81, doi:10.1105/tpc.111.093948.

[7] Alvarez, M.E., Nota, F., & Cambiagno, D.A. “Epigenetic Control of Plant Immunity.” *Molecular Plant Pathology*, vol. 11, no. 4, 1 Feb. 2010, pp. 563–576., doi:10.1111/j.1364-3703.2010.00621.x.

[8] Garcia, H.G., Grayson, P., Han, L., Inamdar, M., Kondev, J., Nelson, P.C., Phillips, R., Widom, J., & Wiggins, P.A. “Biological Consequences of Tightly Bent DNA: The Other Life of a Macromolecular Celebrity.” *Biopolymers*, vol. 85, no. 2, 2007, pp. 115–130., doi:10.1002/bip.20627.

[9] Ceroni, F., & Ellis. T. “The Challenges Facing Synthetic Biology in Eukaryotes.” *Nature Reviews Molecular Cell Biology*, vol. 19, no. 8, 2018, pp. 481–482., doi:10.1038/s41580-018-0013-2.

[10] Medford, J.I., & Prasad,“Plant Synthetic Biology Takes Root.” *Science*, vol. 346, no. 6206, Sept. 2014, pp. 162–163., doi:10.1126/science.1261140.

[11] Medford, J. I., & McCarthy, D. M. (2017). “Growing beyond: Designing Plants to Serve Human and Environmental Interests.” *Current Opinion in Systems Biology*, vol. 5, 2017, pp. 82–85., doi:10.1016/j.coisb.2017.08.008.

[12] Nemhauser, Jennifer L., & Keiko U. Torii. “Plant Synthetic Biology for Molecular Engineering of Signalling and Development.” *Nature Plants*, vol. 2, no. 3, 2 Mar. 2016, doi:10.1038/nplants.2016.10.

[13] Paine, J.A., Shipton, C.A, Chaggar, S., Howells, R.M., Kennedy, M.J., Vernon, G., Wright, S.Y., Hinchliffe, E., Adams, J.L., Silverstone, A.L., & Drake, Rachel. “Improving the Nutritional Value of Golden Rice through Increased Pro-Vitamin A Content.” Nature Biotechnology, vol. 23, no. 4, 2005, pp. 482–87.

[14] Brookes, G., & Barfoot, P. “Economic Impact of GM Crops.” *GM Crops & Food*, vol. 5, no. 1, 1 Jan. 2014, pp. 65–75., doi:10.4161/gmcr.28098.

[15] Shen, B.r., Wang, L.m., Lin, & X., Yao, Z.“Engineering a New Chloroplastic Photorespiratory Bypass to Increase Photosynthetic Efficiency and Productivity in Rice.” *Molecular Plant*, vol. 12, no. 2, 10 Jan. 2019, pp. 199–214., doi:10.1016/j.molp.2018.11.013.

[16] Rocha-Munive, M.G., Soberón, M., Castañeda, S., Niaves, E., Scheinvar,03 E., & Eguiarte, L.E. “Evaluation of the Impact of Genetically Modified Cotton After 20 Years of Cultivation in Mexico.” *Frontiers in Bioengineering and Biotechnology*, vol. 6, 22 June 2018, doi:10.3389/fbioe.2018.00082.

[17] Brophy, J. A. N. & Voigt, C. A. "Principles of genetic circuit design." *Nature Methods* vol. 11, no. 6, 2014, pp. 508–520.

[18]Wolter, F., Klemm, J., & Puchta, H.“Efficient in Planta Gene Targeting in Arabidopsis Using Egg Cell-Specific Expression of the Cas9 Nuclease of Staphylococcus Aureus.” *The Plant Journal*, vol. 94, no. 4, 24 Mar. 2018, pp. 735–746., doi:10.1111/tpj.13893.

[19] Chen, Y., Ho, J., Shis,D., Gupta, C., Long, J., Wagner, D., Ott, W., Josic, K. & Bennett, M. “Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors.” *Nature Communications*, vol. 9, no. 64, 4 Jan. 2018. doi.org/10.1038/s41467-017-02473-5.

[20] T. Gardner, C. Cantor, &J.J. Collins “Construction of a Genetic Toggle Switch in Escherichia Coli.” *Nature News*, Nature Publishing Group, 20 Jan. 2000.

[21] Feng, J., Jester, B.W., Tinberg, C.E., Mandell, D.J., Antunes, M.S., Chari, R., Morey, K.J., Rios, X., Medford, J.I., Church, G.M., Fields, S., & Baker, D. “A General Strategy to Construct Small Molecule Biosensors in Eukaryotes.” *ELife*, vol. 4, 29 Dec. 2015, doi:10.7554/eLife.10606.

[22] Bick, M.J., Greisen, P.J., Morey, K.J., Antunes, M.S., La, D., Sankaran, B., Reymond, L., Johnsson, K., Medfor, J.I., & Baker. D. “Computational Design of Environmental Sensors for the Potent Opioid Fentanyl.” *ELife*, vol. 6, 19 Sept. 2017, doi:10.7554/elife.28909.

[23]McCutcheon, S.R., Chiu, K.L., Lewis, D.D., & Tan, C. “CRISPR-Cas Expands Dynamic Range of Gene Expression From T7RNAP Promoters.” *Biotechnology Journal*, vol. 13, no. 5, 6 Dec. 2017, p. 1700167., doi:10.1002/biot.201700167.

[24] Veylder, L.D., Beeckman, T., Montagu, M.V., & Inzé, D. “Increased Leakiness of the Tetracycline-Inducible Triple-Op Promoter in Dividing Cells Renders It Unsuitable for High Inducible Levels of a Dominant Negative CDC2aAt Gene.” *Journal of Experimental Botany*, vol. 51, no. 351, Jan. 2000, pp. 1647–1653., doi:10.1093/jexbot/51.351.1647.

[25] Tinberg, C.E., Khare, S.D., Dou, J., Doyle, L., Nelson, J.W., Schena, A., Jankowski, W., Kalodimos, C.G., Johnsson, K., Stoddard, B.L., & Baker, D. “Computational Design of Ligand-Binding Proteins with High Affinity and Selectivity.” *Nature*, vol. 501, no. 7466, Sept. 2013, pp. 212–216., doi:10.1038/nature12443.

[26] Ito, H., & Gray, W.M. “A Gain-of-Function Mutation in the Arabidopsis Pleiotropic Drug Resistance Transporter PDR9 Confers Resistance to Auxinic Herbicides.” *Plant Physiology*, vol. 142, no. 1, 28 July 2006, pp. 63–74., doi:10.1104/pp.106.084533.

[27] Ruzicka, K., Strader, L.C., Bailly, A., Yang, H., Blakeslee, J., Langowski, L., Nejedla, E., Fujita, H., Itoh, H., Syono, K., Hejatko, J., Gray, W.M., Martinoia, E., Geisler, M., Bartel, B., Murphy, A.S., & Friml, J.“Arabidopsis PIS1 Encodes the ABCG37 Transporter of Auxinic Compounds Including the Auxin Precursor Indole-3-Butyric Acid.” *Proceedings of the National Academy of Sciences*, vol. 107, no. 23, 24 June 2010, pp. 10749–10753., doi:10.1073/pnas.1005878107.

[26] Schleifenbaum, F., Elgass, K., Sackrow, M., Caesar, K., Berendzen, K., Meixner, A.J., & Harter, K. “Fluorescence Intensity Decay Shape Analysis

Microscopy (FIDSAM) for Quantitative and Sensitive Live-Cell Imaging: A Novel Technique for Fluorescence Microscopy of Endogenously Expressed Fusion-Proteins.” *Molecular Plant*, vol. 3, no. 3, May 2010, pp. 555–562., doi:10.1093/mp/ssp110.

[27] Benfey, P., Bennet, M., & Schiefelbein, J. “Getting to the Root of Plant Biology: Impact of the Arabidopsis Genome Sequence on Root Research.” *The Plant Journal*, vol. 61, no. 6, 1 Mar. 2010, pp. 992–1000., doi:10.1111/j.1365-313x.2010.04129.x.

[28] Roederer, M., & Murphy, R.F. “Cell-by-Cell Autofluorescence Correction for Low Signal-to-Noise Systems: Application to Epidermal Growth Factor Endocytosis by 3T3 Fibroblasts.” *Cytometry*, U.S. National Library of Medicine, Nov. 1986

[29] Bischoff, Wolfgang. “A Functional Central Limit Theorem for Regression Models.” *The Annals of Statistics*, vol. 26, no. 4, Aug. 1998, pp. 1398–1410., doi:10.1214/aos/1024691248.

[30] Otsu, Nobuyuki. “A Threshold Selection Method from Gray-Level Histograms.” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, 1979, pp. 62–66., doi:10.1109/tsmc.1979.4310076.

[31] Pang, Z., Barash, E., Santamaria-Pang, A., Sevinsky, C., Li, Q., & Ginty, F. “Autofluorescence Removal Using a Customized Filter Set.” *Microscopy Research and Technique*, vol. 76, no. 10, 16 July 2013, pp. 1007–1015., doi:10.1002/jemt.22261.

[32] Ito, H., & Gray, W.M. “A Gain-of-Function Mutation in the Arabidopsis Pleiotropic Drug Resistance Transporter PDR9 Confers Resistance to Auxinic Herbicides.” *Plant Physiology*, vol. 142, no. 1, 28 July 2006, pp. 63–74., doi:10.1104/pp.106.084533.

[33] Takeuchi, M., Watanabe, A., Tamura, M., & Tsutsumi, Y. “The Gene Expression Analysis of Arabidopsis Thaliana ABC Transporters by Real-Time PCR for Screening Monolignol-Transporter Candidates.” *Journal of Wood Science*, vol. 64, no. 5, Jan. 2018, pp. 477–484., doi:10.1007/s10086-018-1733-9.

[34] Ruzicka, K., et al. “Arabidopsis PIS1 Encodes the ABCG37 Transporter of Auxinic Compounds Including the Auxin Precursor Indole-3-Butyric Acid.” *Proceedings of the National Academy of Sciences*, vol. 107, no. 23, 2010, pp. 10749–10753., doi:10.1073/pnas.1005878107.

[35] Schürholz, A.k., López-Salmerón, V., Li, Z., Forner, J., Wenzl, C., Gaillochet, C., Augustin, S., Barro A.M, Fuchs, M., Gebert, M., Lohmann, J.U.,

Greb, T., & Wolf, S. “A Comprehensive Toolkit for Inducible, Cell Type-Specific Gene Expression in Arabidopsis.” *Plant Physiology*, vol. 178, no. 1, 2018, pp. 40–53., doi:10.1104/pp.18.00463.

[36] Zuo, Jianru, Niu, Q.w., & Chua, N.h. “An Estrogen Receptor-Based Transactivator XVE Mediates Highly Inducible Gene Expression in Transgenic Plants.” *The Plant Journal*, vol. 24, no. 2, 2000, pp. 265–273., doi:10.1046/j.1365-313x.2000.00868.x.

[36] Roslan, H.A., Salter, M.G., Wood, C.D., White, M.R.H, Croft, K.P., Robson, F., Coupland, G., Doonan, J., Laufs, P., Tomsett, A.B., & Caddick, M.X. “Characterization of the Ethanol-Inducible Alc Gene-Expression System in Arabidopsis Thaliana.” *The Plant Journal*, vol. 28, no. 2, 23 Dec. 2001, pp. 225–235., doi:10.1046/j.1365-313x.2001.01146.x.

[38] Yang, S., Yu, Q., Zhang, Y., Jia, Y., Wan, S., Kong, X., & Ding, Z. “ROS: The Fine-Tuner of Plant Stem Cell Fate.” *Trends in Plant Science*, vol. 23, no. 10, 2018, pp. 850–853., doi:10.1016/j.tplants.2018.07.010.

[39] Pěnčík, A., Simonovik, B., Petersson, S.V., Henykova, E., Simon, S., Greenham, K., Zhang, Y., Kowalcyzk, M., Estelle, M., Zažímalová, E., Novak, O., Sandberg, G., & Ljung, K. “Regulation of Auxin Homeostasis and Gradients in Arabidopsis Roots through the Formation of the Indole-3-Acetic Acid Catabolite 2-Oxindole-3-Acetic Acid.” *The Plant Cell*, vol. 25, no. 10, 2013, pp. 3858–3870., doi:10.1105/tpc.113.114421.

[40] Fu, Z.Q., Yan, S., Saleh, A., Wang, W., Ruble, J., Oka, N., Mohan, R., Spoel, S.H., Tada, Y., Zheng, N., & Dong, X. “NPR3 And NPR4 Are Receptors for the Immune Signal Salicylic Acid in Plants.” *Nature*, vol. 486, no. 7402, 2012, pp. 228–232., doi:10.1038/nature11162.

[41] Kang, H.g, Fang, Y., & Singh, K.B. “A Glucocorticoid-Inducible Transcription System Causes Severe Growth Defects in Arabidopsis and Induces Defense-Related Genes.” *The Plant Journal*, vol. 20, no. 1, Oct., 1999, pp. 127–133., doi:10.1046/j.1365-313x.1999.00575.x.

[42]Aoyama, T., & Chua, N.H. “A Glucocorticoid-Mediated Transcriptional Induction System in Transgenic Plants.” *The Plant Journal*, vol. 11, no. 3, 1997, pp. 605–612., doi:10.1046/j.1365-313x.1997.11030605.x.

[43] Vassilios, S., & Kaznessis, Y.N. “Synthetic Tetracycline-Inducible Regulatory Networks: Computer-Aided Design of Dynamic Phenotypes.” *BMC Systems Biology*, vol. 1, no. 1, Sept. 2007, doi:10.1186/1752-0509-1-7.

[44] Rong, G., & Stock, A.M.. “Temporal Hierarchy of Gene Expression Mediated by Transcription Factor Binding Affinity and Activation Dynamics.” *MBio*, vol. 6, no. 3, 2015, doi:10.1128/mbio.00686-15.

[45] Shaunak K,, & Ellington, A.D. “Construction of Synthetic T7 RNA Polymerase Expression Systems.” *Methods*, vol. 143, 1 July 2018, pp. 110–120., doi:10.1016/j.ymeth.2018.02.022.

[46] Covington, M.F., & Harmer, S.L. “The Circadian Clock Regulates Auxin Signaling and Responses in Arabidopsis.” *PLoS Biology*, vol. 5, no. 8, July 2007, doi:10.1371/journal.pbio.0050222.

[47] Ke, M., Gao, Z., Chen, J., Qiu, Y., Zhang, L., & Chen, X. “Auxin Controls Circadian Flower Opening and Closure in the Waterlily.” *BMC Plant Biology*, vol. 18, no. 1, 11 July 2018, doi:10.1186/s12870-018-1357-7

[48] Li, Y., Wang, H., Zhang, Y., & Martin, C. “Can the World’s Favorite Fruit, Tomato, Provide an Effective Biosynthetic Chassis for High-Value Metabolites?” *Plant Cell Reports*, vol. 37, no. 10, 28 Mar. 2018, pp. 1443–1450., doi:10.1007/s00299-018-2283-8.

Chapter 3: Incorporation of a positive feedback motif into an inducible, root epidermal cell-specific genetic circuit

3.1 Introduction

Multicellular organisms evolved regulatory mechanisms and networks to divide labor across cells and tissues composing whole organisms. Part of these mechanisms are regulatory genetic motifs. For instance, transcriptional repressors and activators with feedback between each other have been implicated in pattern-triggered immunity in plants, floral organ abscission, and circadian clocks [1,2,3]. Additionally, regulatory motifs have been encoded into synthetic genetic circuits to achieve desired behaviors and outputs as a proof-of-concept [4,5,6]. Specifically, autoregulatory positive feedback loops have been successfully implemented in bacterial, yeast, and mammalian cells to amplify output, modify circuit kinetics, be more sensitive to input, and impart cellular memory or hysteresis. A commonality between these systems is an ultrasensitive response, wherein a small increase of input results in a drastic increase of the total output [7,8]. This drastic increase in output can approximate a digital or switch-like response.

The ability to design and incorporate positive feedback into synthetic circuits functioning within plant platforms could be extremely valuable. For instance, plants could be engineered to sense the level and persistence of various environmental toxins and respond by initiating a transcription regulatory network, after a certain threshold of the toxin is met, to remediate the soil [9]. Even with a reduction of the toxin, the plant should

have memory of the toxin and continue to initiate the phytoremediation pathway. Additionally, plants could be designed such that biofuel production is switched-on after a specific developmental stage is reached [10]. To this end, we developed a positive feedback genetic system functioning within *Arabidopsis thaliana.*

The inducible, root epidermal cell-specific genetic circuit detailed in the previous chapter was expanded upon with a simple positive feedback motif. In this design, the positive feedback is dependent on a transcription factor (Gal4-VP64/PF) driving transcription of itself (Figure 3.1). Transcription of the PF transcription factor was driven by the inducible (UAS) promoter. Driving transcription of the PF transcription factor with the UAS promoter maintained response to digoxin, therefore the genetic circuit would still be inducible and quantitatively controlled with input. Encoding the inducible promoter required that the PF transcription factor be able to bind the UAS promoter, therefore it has a Gal-4 DNA-binding domain (DBD). Preliminary results showed the VP64 activation domain resulted in the best activation and was used to partially compose the PF transcription factor.

Demonstrated in this chapter is a simple positive feedback genetic circuit that has targeted output to root-epidermal cells. Additionally the circuit is inducible and more sensitive to input than the comparable genetic circuit detailed in chapter two. Furthermore, the positive feedback genetic circuit responses faster to ligand than the circuit and has what appears to be memory when transiently exposed to ligand.

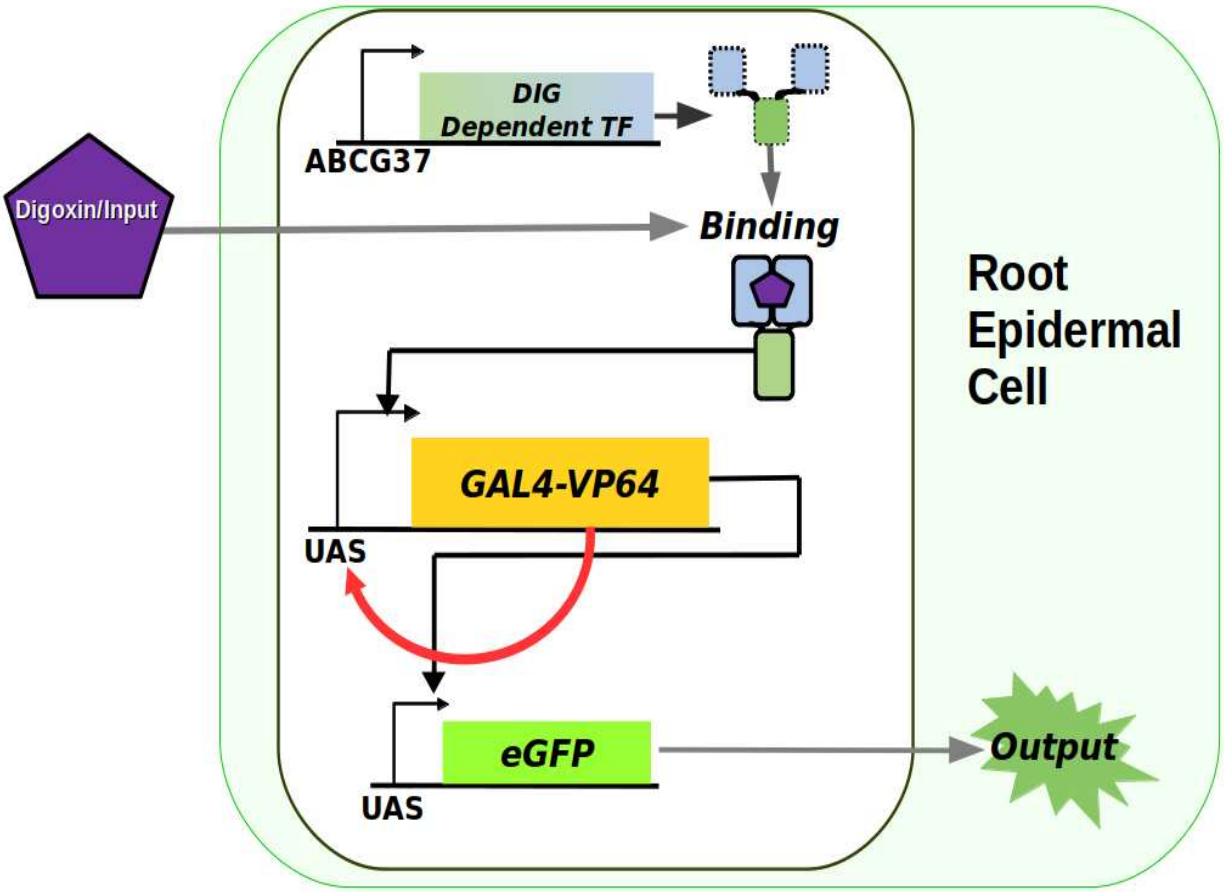


Figure 3.1. Topology of the positive feedback genetic circuit. The ABCG37 promoter drives transcription of a computationally designed, DIG dependent TF. The DIG dependent TF can bind the inducible promoter (UAS). The UAS promoter drives the expression of the reporter molecule, enhanced green fluorescent protein (GFP). The DIG dependent TF "jump starts" the positive feedback by binding the UAS promoter driving expression of the GAL4-VP64 (positive feedback/PF) TF. The transcription factor is autoregulatory, *i.e.*, can drive transcription of itself after translation and expression of the GAL4-VP64 transcription factor. Moreover the PF transcription factor can drive transcription of GFP.

3.2 Methods

3.2.1 Assembly of positive feedback genetic circuit

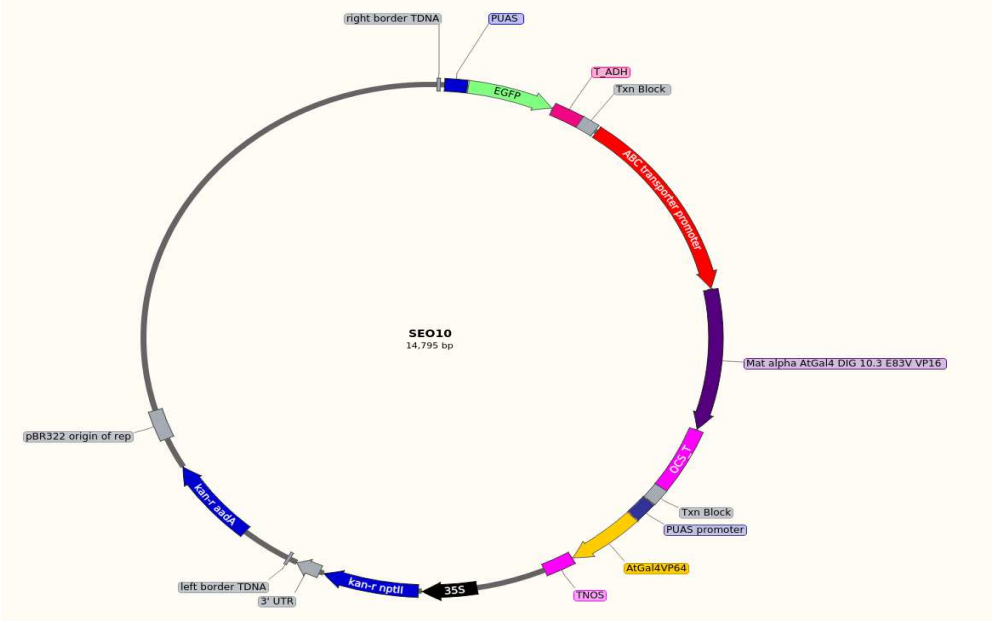
The ABCG37 promoter was amplified using KJM325 vector (Table 3.1) as template with Phusion high fidelity polymerase (NEB, Ipswitch, MA) and primers 7 and 8 (Table A.4) to add KpnI and PacI restriction sites to the 5’ and 3’ ends of the amplicon, respectively. The PCR product was screened as in chapter two. The ABCG37 promoter was directionally cloned into KpnI and PacI digested KJM412-pCambia 2300 binary vector (Table 3.2) using T4 ligase (NEB, Ipswitch, MA) upstream of the ligand dependent transcription factor to drive expression of it and impart tissue specificity to the circuit. The resulting plasmid is KJM414 (Table 3.2). The inducible promoter (UAS promoter) and reporter (GFP) with the terminator and transcription block, in addition to the 5’ end of the ABC were removed from SEO4 plasmid using KpnI and PacI restriction enzymes (NEB, Ipswitch, MA). This DNA fragment was directionally cloned into KpnI and PacI digested KJM414 using T4 ligase downstream of the left T-DNA boarder and upstream of the ABCG37 promoter to produce SEO10 (Figure 3.2).

Table 3.1: Template for PCR

Construct Name	Description (5’→3’)	Forward Primer	Reverse Primer	Template sequence/PCR product	Added Restriction Sites (5’, 3’)
KJM325	ABCG37 promoter::Myb41::TNOS::Txn block	7	9	UAS promoter	Kpn, PacI

Table 3.2: Constructs for cloning SEO10 plasmid

Purpose	Construct Name	Description (5'→3')	Restriction enzymes digested with	Removed fragments	Cloned fragments	Resulting construct from cloning
Sub clone of ABCG37 promoter	KJM 412	35S promoter::DIG dependent TF::OCS::Txn. UAS promoter::Gal4-VP64-TNOS::Txn block	KpnI, PacI	35S	UAS promoter	KJM414
Source of PF and DIG systems	KJM 414	ABCG37 promoter::DIG dependent TF::OCS::Txn. UAS promoter::Gal4-VP64-TNOS::Txn block	KpnI, PacI	N/A	UAS::GFP::	SEO10
Source of UAS, GFP and terminator for final construct	SEO 4	ABCG37 promoter::DIG dependent TF::TNOS::Txn block	KpnI, PacI	UAS::GFP::T-ADH::TXN block. 5' fragment of ABCG37 promoter	N/A	SEO10



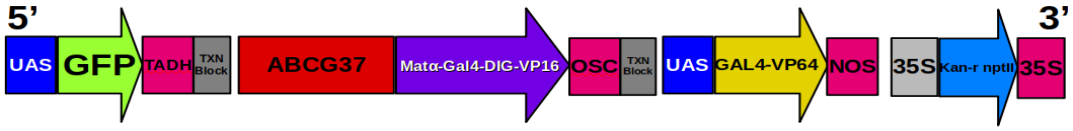


Figure 3.2. Construct (SEO10) and linearized T-DNA of positive feedback circuit. The circuit is composed of three genes. From the left t-DNA boarder, is a UAS promoter driving the transcription of eGFP, followed by an TADH. To ensure read through does not occur a transcription block proceeds the TADH sequence. The subsequent DNA unit is composed the ABCG37 promoter driving the transcription of the DIG dependent TF and terminated with a OCS terminator. The next DNA unit is a UAS promoter driving transcription of Gal4-VP64 transcription factor and terminated with a NOS terminator. Kanamycin markers were used for selection. Transcription of the kanamycin marker is constitutive and driven by a 35S CaMV promoter and terminated by 35S terminator. Figure generated using SnapGene software.

3.2.3 Screening bacterial, and T0 and T1 transformants

For general methods performed to identify stably transformed Arabidopsis plants with the positive feedback genetic circuit reference chapter 2, section 2.2.2. SEO10 *E.coli* transformants were screened using colony PCR with primers 34 and 35 and SEO1 plasmid as a positive control.

Three positive *E.coli* colonies were identified. Plasmids were isolated as in chapter two and sent to Macrogen (Rockville, MD).Samples with approximately 200 ng/μL of SEO10 plasmid and one of the primers from primers nine to 34 were sent in as reactions for sequencing (A.4).

Chapter two also details ,in section 2.2.3, methods to screen T0 and T1 plants for inducible and root epidermal specific GFP signal. Only two transformants of the total eleven screened in the T1 generation were positive for GFP. This is unlike chapter two, wherein all T1 transfomants segregating three to one for kanamycin resistance were positive for GFP signal after induction. Additionally, RNA for GFP was isolated using a easy-spin IIp Plant RNA Extraction Kit (Qiagen, Hilden, DE). The extracted RNA was used as template for RT-PCR with primers one and two and Phusion DNAP. The cDNA was run on a 1% agarose electrophoresis gel to screen for presence of bands the size of the GFP sequence within the positive feedback transformants.

3.2.4 Confocal microscopy to quantify GFP signal

Confocal microscopy approaches for positive feedback roots were performed in the same manner as in chapter two. However, only plants from the T1 generation were used for confocal microscopy. Moreover, only one transgenic line was responsive to positive for GFP and inducible, therefore only this line was analyzed. Additionally, preliminary results suggested the two transgenic lines had observable GFP signal after four hours of continuous induction, therefore temporal characterization experiments were shifted to begin screening after four hours and did not include a six hour time point.

3.2.5 Comparing roots with and without positive feedback

SEO1-50 and SEO10-225 were used for comparison between genetic circuits with and without positive feedback, respectively. These lines were chosen because each had the largest dynamic range of the lines analyzed with and without positive feedback. Seed of T2 (SEO1-50) and T1 (SEO10-225) lines, and wild-type plants were sterilized and plated on vertical MS media plates containing kanamycin (50µl/ml) antibiotic. After 10 to 12 days of growth under a light cycle of 16 hours of light and 8 hours of darkness, SEO1-50, SEO10-225, and wild-type plants were transferred to 24 well plates containing liquid MS media in the absence or presence of various digoxin concentrations including 10nM, 100nM, 1µM , 10µM, 100µM. After 16 hours of continuous induction, plants were washed in fresh liquid MS media in a clean 24-well plate. Plates were placed on a shaker set at 50 rpm for an hour. Two washes were performed. After washing, the roots were harvested. Harvested roots were mounted on glass slides in 50% glycerol and imaged using a Olympus IX83 Inverted Spinning Disk Confocal Microscope with a 488nm excitation laser, and 525/30nm and 617/73nm emission filters. All samples were exposed to the 488nm laser for 500ms. Z stacks were taken at each interrogated point over 80 microns. The resulting images were run through the MATLAB program detailed in chapter two section 2.2.5.

3.3 Results

3.3.1 Root epidermal specificity and positive feedback

GFP signal was extremely evident in the root epidermis (Figure 3.3). This included extensions of the epidermis (root hairs). Furthermore, the same spatial pattern dependent observed without positive feedback is observed in transgenic epidermal cells containing the positive feedback genetic circuit. Lateral roots and the zone of elongation GFP signal were greater and lesser than all other zones, respectively (Figure 3.4). Specifically, the lateral samples had a GFP signal intensity three fold greater than the zone of elongation samples. Moreover, the zones of maturation and division had GFP signal greater than the zone of elongation, but less than the lateral roots. The lateral root GFP signal was 37%, 32%, and 51% greater than the zones of division, proximal and distal zone of maturation, respectively. The zone of elongation had a two fold decrease in GFP signal to the distal maturation and division zones. The GFP signal of the proximal zone of maturation was nearly 30% greater than that of the zone of elongation. Additionally, the zone of division had a greater GFP signal than the proximal portion of the zone of maturation by 30%, but was nearly the same as the distal portion. The spatial pattern of GFP signal was highly similar to those observed without positive feedback (Figure 2.9, 2.10, 2.11). For both genetic circuits the most intense and least intense signals were observed in the lateral roots and zone of elongation, respectively.

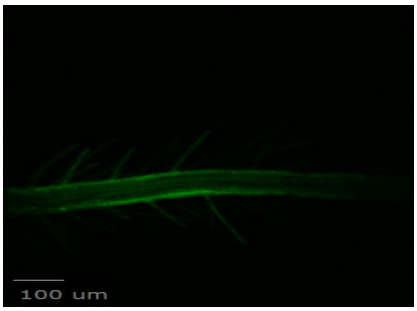


Figure 3.3. Confocal image of root harvested from positive feedback T1 transfomant SEO10-225. Representative transgenic root continually induced with 1 µM DIG liquid media for 16 hours.

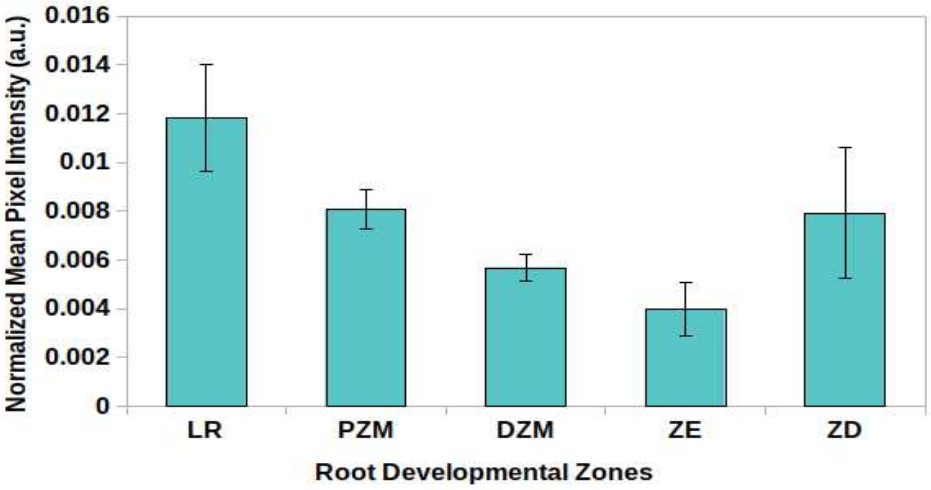


Figure 3.4. Expression of GFP is under developmental regulation due to root-epidermal specific (ABCG37) promoter. Root samples were treated for 16 hours. From left to right: ZD: zone of division, ZE: zone of elongation, PZM: proximal zone of maturation, DZM: distal zone of maturation. LR: lateral roots. Confocal images were normalized to transgenic controls to account for basal level of expression .n=6 per treatment. Error bars represent +/- 2 SE.

3.3.2: Positive feedback genetic circuit over digoxin concentrations

The GFP signal of the transgenic lines differed. SEO10-225 at 1, 10, and 100 µM ligand concentrations was significantly greater than transgenic controls (Figure 3.5). However, the only significantly different GFP signal in

SEO10-240 was observed at a ligand concentration 1 μM. Moreover, the fold induction of this line over transgenic controls was under two fold for all concentrations (Figure 3.6). Furthermore, when the standard deviation error term was applied to the output a significant difference did not exist between transgenic controls and those treated with 1 μM ligand. The other transgenic line had a nearly 30, 19, and 14 at ligand concentrations of 1, 10, and 100μM, respectively. Furthermore, the response curve of this line appears to be sigmoidal in shape (Figure 3.5). Concentrations below 1 μM did not produce GFP signal significantly different than the transgenic control, a sharp increase occurs at a 1 μM ligand concentration, and samples under ligand concentrations above 1 μM have a GFP signal.

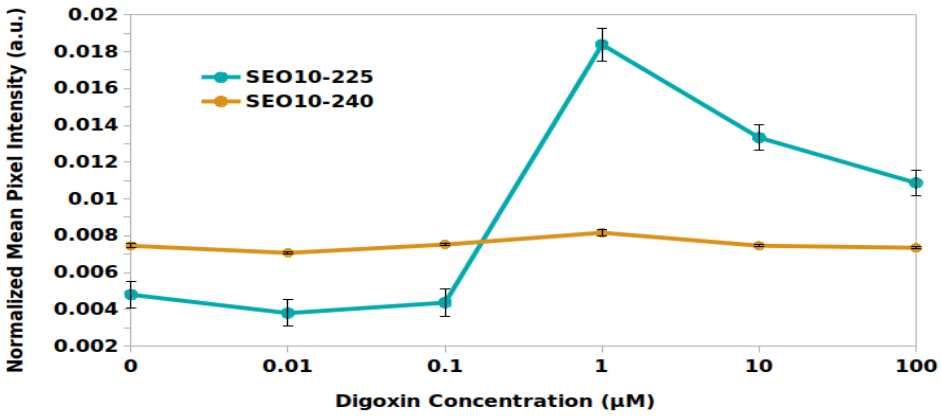


Figure 3.5. Positive feedback genetic circuit over digoxin concentrations. Zero concentration represents the transgenic control. Both transgenic lines were continually induced for 16 hours at each treatment concentration. Samples have been normalized to wild-type samples. n=4 per treatment, three technical replicates per sample. Error bars represent +/- 2 SE.

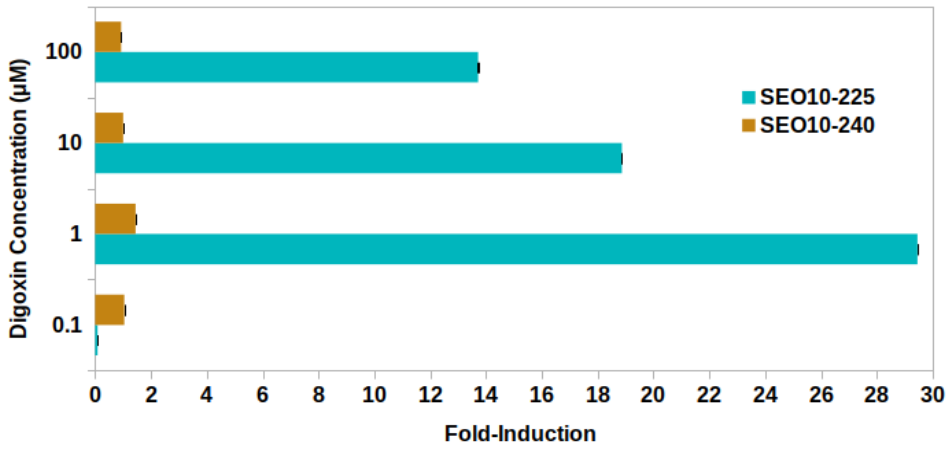


Figure 3.6. Fold-induction with incorporated positive-feedback loop. Fold induction of each SEO10 independent transgenic lines. Fold induction was calculated by normalizing to wild-type sample and calculating the ratio of GFP signal between induced transgenic samples and transgenic controls. n=4 per treatment, three technical replicates per sample. Error bars represent the 95% confidence interval.

3.3.3: Activation characterization of positive feedback circuit

The positive feedback genetic circuit had significantly different GFP signal above transgenic controls after four hours of continuous induction (Figure 3.7). The positive feedback exhibited a faster activation response relative to the inducible tissue specific genetic circuit lacking the positive feedback by two hours (Figure 2.13). However, both genetic circuits exhibited maximal GFP signal after 16 hours of continuous induction and a dramatic decrease in expression after 24 hours of continuous induction. The GFP signal in treated samples after 16 hours of continual induction was over 15 fold greater than transgenic samples continually induced for 24 hours. Furthermore, the GFP signal after four hours of continuous induction was also greater than the GFP signal after 24 hours of continuous induction.

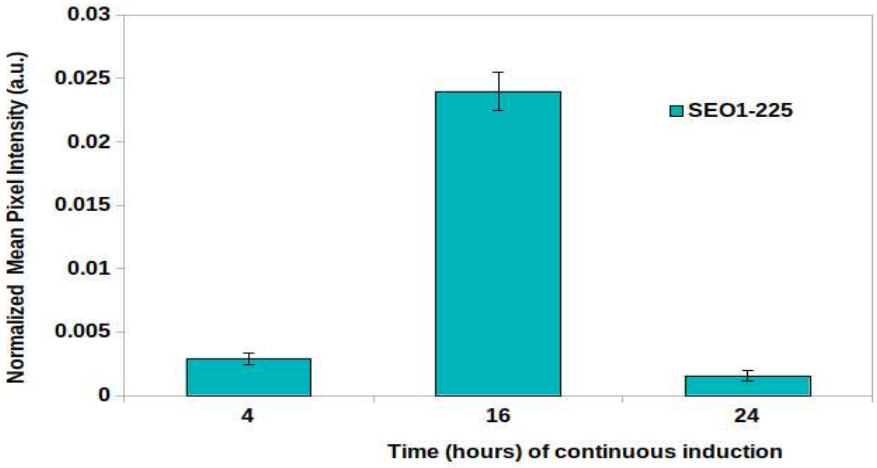


Figure 3.7. Positive feedback time course. Treated samples were continually induced for 16 hours. Treatment group samples were subjected to liquid MS media containing 1μM ligand. Images of treatment group outputs have been normalized to transgenic controls at each time point to account for basal expression. n=4 per treatment, three technical replicates per sample. Error bars represent +/- 2 SE.

3.3.4: Positive feedback: Sensitivity and amplification

The transgenic line with the genetic circuit containing the positive feedback and responsive to ligand, exhibited maximal GFP signal after being elicited with 1 μM ligand (Figure 3.5, 3.8). Transgenic lines with the genetic circuit without positive feedback exhibited maximal output with ligand treatment of 100 μM (Figure 2.16, 3.8). In other words, the sensitivity of the genetic circuit with the positive feedback loop motif was 100 fold more than that of the genetic circuit without this motif. However, the maximal output of the genetic circuit without positive feedback was greater than that of the genetic circuit with positive feedback by four fold (Figure 3.8).

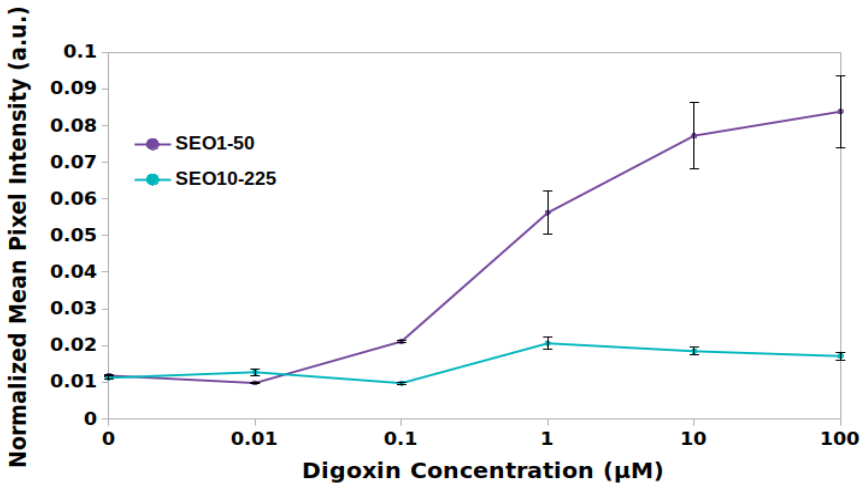


Figure 3.8. Amplification and sensitivity: Comparison between positive feedback and non-positive feedback genetic circuits. All samples were treated for 16 hours. Induced transgenic samples and transgenic controls have been normalized to wild-type samples. n= 4 per line, three technical replicates per sample. Error bars represent +/- 2 SE. Transgenic controls were subjected to media containing carrier (DMSO) for DIG.

3.3.5 Memory and positive feedback

After one day from the removal of transgenic samples from induction media output decreased to nearly 50% of that observed in continually induced samples (Figure 3.9). However, output reached what appeared to be a steady level and remained at this level for a period of at least three days up to five days. Transgenic controls had GFP signal greater than those screened one and three days after removal from induction media. Moreover, outputs of transgenic control samples were greater than those samples induced five days prior. Due to this response, it is difficult to identify if GFP signal is retained up to five days after removal from induction or the source of increased GFP signal.

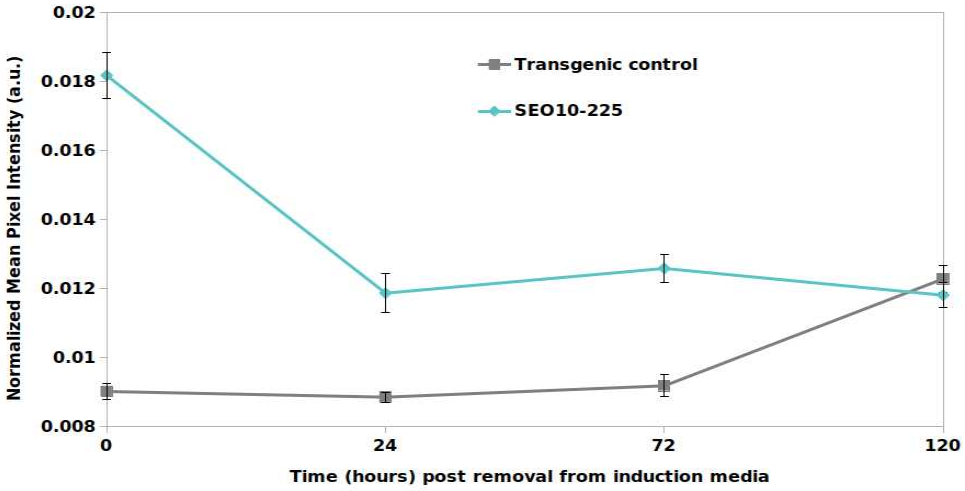


Figure 3.9. Positive feedback motif may impart limited memory
All samples were treated (absence of presence of ligand) continually for 16 hours. Zero time point represents those samples remove from treatment screened immediately after removal from induction media. n= 4 per line, three technical replicates per sample for each time point. Error bars represent +/- 2 SE. Transgenic controls were subjected to medium containing carrier (DMSO) for DIG.

3.4 Discussion

Incorporation of a simple positive feedback motif into the root epidermal cell-specific DIG transactivation system modifies the dynamics and sensitivity of the genetic system. However, incorporating positive feedback did not disrupt inducible and spatial expression patterns observed without positive feedback. Moreover, the spatial GFP signal patterns of the root epidermal specific DIG transactivation system with and without positive feedback appear to be the same (Figure 2.11, 3.4). That is, GFP signal occurred solely within the root epidermal tissue and along the entirety of roots. Moreover, GFP signal followed the developmental zones of the root.

Specifically, the overall GFP signal of lateral roots and the zone of elongation were greater and lesser than all other developmental zones, respectively (Figure 3.4). For instance, the lateral roots had 30 to 50% greater GFP signal than the zones of division and maturation, respectively, and two fold greater signal than the zone of elongation. GFP signal within the zones of maturation and division were two fold greater than the GFP signal within the zone of elongation. These results, as was the case with the genetic circuit without positive feedback, align with studies characterizing the ABCG37 promoter and PDR9 transporter localization [11][12][13]. Hence, I conclude the regulation and characteristics of a promoter driving transcription of the DIG dependent TF will extend to the output of the DIG transactivation system.

The positive feedback incorporated into the inducible, tissue-specific genetic circuit increased sensitivity 100 fold over the genetic circuit without the motif. However, of the two lines analyzed, one (SEO10-240) was not as responsive to ligand and appeared to be in an ON state, *i.e.,* a monostable not bistable response. Other positive feedback systems have had multiple steady states with only one stable steady state [14]. This could be the case for this transgenic line.

The overall output, dynamic range and fold induction of the positive feedback genetic circuit was reduced relative to the genetic circuit with no positive feedback. A possible reason for the lack in amplification could be duplicate UAS promoters and homology between the positive feedback and DIG dependent transcription factor nucleotide sequences. The homology

could have initiated silencing mechanisms at the transcription or post-transcription levels [15][16][17]. However, the homology was not to such a degree for this to be likely. The responsive transgenic line should be further analyzed using RT-qPCR and pull-down assays to test if silencing is occurring at the transcriptional and post-transcriptional level, respectively. Additionally, identical and duplicate promoters utilize the same TF pool possibly imposing competition between UAS promoters driving transcription of GFP and positive feedback. In other systems, mathematical modeling and empirical data have shown duplicate identical promoters within a genetic system impose competition for the same TF pool between promoters, and as a result output decreases [18][19]. Furthermore, positional effects could have influenced a relative decrease in output between the positive and non-positive feedback genetic circuits [17]. Generating more transgenic lines could reveal if positional effects are driving the difference in level of output, particularly if output varies significantly between transgenic lines, and if some lines do have an amplified output relative to non-positive feedback lines.

Although amplification did not occur with the positive feedback genetic circuit it does appear to have an ultrasensitive response. The percent in total GFP signal reached at the next lowest ligand concentrations from the ligand concentrations eliciting the maximum output, in each respective line, was lower in the responsive transgenic line containing positive feedback compared to the transgenic line without positive feedback. Approximately 5 and 91% of the total output was reached at ligand concentrations of 0.1 μM

(positive feedback line) and 10 μM (non-positive feedback lines), respectively. This means 95 and 9% of the response in the positive feedback and non-positive feedback lines, respectively, were achieved over the same 10 fold change in ligand concentration. This represents an ultrasensitive response because a greater percent of the overall output was reached over the same change in input [7][8][14]. An ultrasensitive response can consequently result in bistability and hysteresis.

The positive feedback genetic circuit appears to have memory. An ultrasensitive response paired with low and high states are characteristics of bistable circuits [14]. Furthermore, input and output dose-response curves of the responsive transgenic line were sigmodial in nature and “switch-like,” which is a graphical characteristic of bistable circuits. One of the characteristics of ultrasensitive and bistable responses is memory [4][14]. The positive feedback line output decreases 24 hours after removal from induction media, but after this initial decline, GFP signal is maintained up to five days from removal. However, transgenic control samples had GFP signal increases to a greater level on the fifth day than previous days. A source of this increase is most likely due to experimental artifact. Dimethyl sulfoxide (DMSO) has been shown to cause stress responses and cytotoxicity in *Arabidopsis* [20]. Residual DMSO may have caused an increase in GFP signal. It is more likely to be experimental artifact, given an increase in GFP signal was also observed on the fifth day in transgenic lines without positive feedback (Figure 2.15). Although memory is not clearly defined on the fifth

day, memory can be observed up to and including three days after removal from induction media. However, an initial decline does occur 24 hours from removal. The initial decrease could arise from a circadian rhythm. That is, GFP signal decreased after 24 hours of continual induction in both transgenic roots with and without positive feedback. Roots were observed at 24 hour time points, therefore GFP signal may have declined and increased again after observation. Moreover, Medford lab members have also observed what appeared to be a circadian rhythm with a similar positive feedback design. Nonetheless, transgenic lines without positive feedback did not maintain the output at the same level after the initial decrease over the same time, but continually decreased until GFP signal did not significantly differ between induced samples and transgenic controls on the fifth day. In these transgenic lines, the output decreased by approximately 50% to 70% from the first to the third days after removal from induction media. The positive feedback system can sustain the response after the initial decrease in output, therefore this could be a form of memory.

Interestingly, positive feedback decreased the activation time. Many positive feedback genetic circuits are characterized by a delay in activation relative to comparable circuits without positive feedback [21][22]. However, differences in kinetics between systems and constituent components of those systems consequently create characteristically and dynamically distinct positive feedback systems [22]. Positive feedback systems considered to be weak have had accelerated response times and reduced

amplification. Although the activation time is decreased with positive feedback, the time required to reach the maximal response is equivalent between the positive and non-positive feedback genetic circuits. In this respect, the positive feedback system is delayed. The time required to activate the system and maximal response is two hours longer than the non-positive feedback circuit. This may be the characteristic delay observed in other positive feedback systems.

The genetic circuit may require redesign to achieve the desired properties including amplification and robust memory. A means to create such a design could include *in silico* approaches, namely mathematical modeling, with empirical data concerning each component of the circuit to inform the mathematical model. Specifically, quantitatively defining the input and output of the ABCG37, *i.e.*, defining the transfer function. Mathematical modeling performed by members of the Medford and Prasad labs have shown balancing the three promoter strengths within the positive feedback genetic circuit is crucial to achieve the desire responses. More than likely, the ABCG37 promoter was not balanced within the positive feedback system. The design, particularly the promoter driving transcription of the DIG dependent transcription factor, may need to be revised. Specifically, promoters that have been well characterized should be incorporated into the design. Moreover, promoters that are mathematically balanced. That is a promoter that is stronger than the UAS promoter, and the strongest within the genetic circuit. Several systems have been built with either positive or

negative feedback, or both motifs to decrease activation time, resist and buffer against noise, and create robust switches with memory [22][23][24][25]. The positive feedback system detailed here could possibly achieve these behaviors, or possibly be a simple component of a redesigned higher-order genetic system.

REFERENCES

[1] Li, B., Meng, X., Shan, L., & He, P. “Transcriptional Regulation of Pattern-Triggered Immunity in Plants.” *Cell Host & Microbe*, vol. 19, no. 5, 11 May 2016, pp. 641–650., doi:10.1016/j.chom.2016.04.011.

[2]Patharkar, O Rahul, & John C Walker. “Floral Organ Abscission Is Regulated by a Positive Feedback Loop.” *Proceedings of the National Academy of Sciences of the United States of America*, National Academy of Sciences, 3 Mar. 2015

[3]Drapek, C.E., Sparks, E.E., Benfey, P.N. “Uncovering Gene Regulatory Networks Controlling Plant Cell Differentiation.” *Trends in Genetics*, vol. 33, no. 8, Aug. 2017, pp. 529–539., doi:10.1016/j.tig.2017.05.002.

[4] Ajo-franklin, C.M., Drubin, D.A., Eskin, J.A., Gee, E.P.s., Landgraf, D., Phillips, I., & Silver, P.A. “Rational Design of Memory in Eukaryotic Cells.” *Genes Dev*.-2007-Ajo-Franklin-2271-6. no. 617, 2007, pp. 2271–76, doi:10.1101/gad.1586107.GENES.

[5] Siciliano, V., Menolascina, F., Marucci, L., Fracassi, C., Garzilli, I., Moretti, M.N., Bernardo, & D.D. “Construction and Modelling of an Inducible Positive Feedback Loop Stably Integrated in a Mammalian Cell-Line.” *PLoS Computational Biology*, vol. 7, no. 6, 2011, doi:10.1371/journal.pcbi.1002074.

[6] Nistala, G.J., Wu, K., Rao, C., & Bhaelero, K.D. “A Modular Positive Feedback-Based Gene Amplifier.” *Journal of Biological Engineering*, vol. 4, no. 1, 2010, p. 4., doi:10.1186/1754-1611-4-4.

[7] Goldbeter, A, & D E Koshland. “An Amplified Sensitivity Arising from Covalent Modification in Biological Systems.” *Proceedings of the National Academy of Sciences of the United States of America*, U.S. National Library of Medicine, Nov. 1981.

[8] Goldbeter, A., & Koshland, D.E. “Ultrasensitivity in Biochemical Systems Controlled by Covalent Modification.” The Journal of Biological Chemistry, vol. 259, no. 23, 1984, pp. 14441–47.

[9] Adiloğlu, Sevinç. “Heavy Metal Removal with Phytoremediation.” *Advances in Bioremediation and Phytoremediation*, 4 Dec. 2018, doi:10.5772/intechopen.70330.

[10] Medford, J. I., & McCarthy, D. M. (2017). “Growing beyond: Designing Plants to Serve Human and Environmental Interests.” *Current Opinion in Systems Biology*, vol. 5, 2017, pp. 82–85., doi:10.1016/j.coisb.2017.08.008.

[11] Ito, H., & Gray, W.M. “A Gain-of-Function Mutation in the Arabidopsis Pleiotropic Drug Resistance Transporter PDR9 Confers Resistance to Auxinic Herbicides.” *Plant Physiology*, vol. 142, no. 1, 28 July 2006, pp. 63–74., doi:10.1104/pp.106.084533.

[12] Takeuchi, Manami, et al. “The Gene Expression Analysis of Arabidopsis Thaliana ABC Transporters by Real-Time PCR for Screening Monolignol-Transporter Candidates.” *Journal of Wood Science*, vol. 64, no. 5, Jan. 2018, pp. 477–484., doi:10.1007/s10086-018-1733-9.

[13] Ruzicka, K., et al. “Arabidopsis PIS1 Encodes the ABCG37 Transporter of Auxinic Compounds Including the Auxin Precursor Indole-3-Butyric Acid.” *Proceedings of the National Academy of Sciences*, vol. 107, no. 23, 2010, pp. 10749–10753., doi:10.1073/pnas.1005878107.

[14] Ferrell, James E., and Sang Hoon Ha. “Ultrasensitivity Part III: Cascades, Bistable Switches, and Oscillators.” *Trends in Biochemical Sciences*, vol. 39, no. 12, 14 Nov. 2014, pp. 612–618., doi:10.1016/j.tibs.2014.10.002.

[15] Carbonell, A. “Artificial Small RNA-Based Strategies for Effective and Specific Gene Silencing in Plants.” *Plant Gene Silencing: Mechanisms and Applications*, June 2000, pp. 110–127., doi:10.1079/9781780647678.0110.

[16] Brewster, Robert C., et al. “The Transcription Factor Titration Effect Dictates Level of Gene Expression.” *Cell*, vol. 156, no. 6, 13 Mar. 2014, pp. 1312–1323., doi:10.1016/j.cell.2014.02.022.

[17] Elomaa, Paula, et al. “Transgene Inactivation In Petunia Hybrida Is Influenced by the Properties of the Foreign Gene.” *Molecular and General Genetics MGG*, vol. 248, no. 6, 25 Oct. 1995, pp. 649–656., doi:10.1007/bf02191704.

[18] Rydenfelt, Mattias, et al. “Statistical Mechanical Model of Coupled Transcription from Multiple Promoters Due to Transcription Factor Titration.” *Physical Review E*, vol. 89, no. 1, 6 Jan. 2014, doi:10.1103/physreve.89.012702.

[19] Holtz, William J., and Jay D. Keasling. “Engineering Static and Dynamic Control of Synthetic Pathways.” *Cell*, vol. 140, no. 1, 8 Jan. 2010, pp. 19–23., doi:10.1016/j.cell.2009.12.029.

[20] Shibasaki K., Uemura, M., Tsurumi, S., & Rahman, A.“Auxin Response in Arabidopsis under Cold Stress: Underlying Molecular Mechanisms.” *The Plant Cell*, vol. 21, no. 12, 9 Dec. 2009, pp. 3823–3838., doi:10.1105/tpc.109.069906.

[21] Brandman, O. “Interlinked Fast and Slow Positive Feedback Loops Drive Reliable Cell Decisions.” *Science*, vol. 310, no. 5747, 21 Oct. 2005, pp. 496–498., doi:10.1126/science.1113834.

[22] Chang, D.-E., et al. “Building Biological Memory by Linking Positive Feedback Loops.” *Proceedings of the National Academy of Sciences*, vol. 107, no. 1, 29 Dec. 2009, pp. 175–180., doi:10.1073/pnas.0908314107.

[23] Holtz, William J., and Jay D. Keasling. “Engineering Static and Dynamic Control of Synthetic Pathways.” *Cell*, vol. 140, no. 1, 8 Jan. 2010, pp. 19–23., doi:10.1016/j.cell.2009.12.029.

[24] Chang, D.-E., et al. “Building Biological Memory by Linking Positive Feedback Loops.” *Proceedings of the National Academy of Sciences*, vol. 107, no. 1, 29 Dec. 2009, pp. 175–180., doi:10.1073/pnas.0908314107.

[25] Ananthasubramaniam, Bharath, and Hanspeter Herzel. “Positive Feedback Promotes Oscillations in Negative Feedback Loops.” *PLoS ONE*, vol. 9, no. 8, 15 Aug. 2014, doi:10.1371/journal.pone.0104761.

Chapter 4: Conclusions and Future works

4.1 Dual-wavelength emission and high-throughput analysis

Analyzing confocal images in a high-throughput manner can be accomplished by processing dual-wavelength emission images through the MATLAB program described in chapter two. This approach could be further tested with a range of fluorescence reporters and emission filters to identify what experimental designs the approach could be applied to. The genetic circuits detailed here are modular, for this reason, GFP could readily be swapped with other types of fluorescence reporters. Furthermore, multiple fluorescence reporters could be quantified simultaneously in one image. Consequently, parts and outputs could be compared without introducing confounding variables arising from experimental variability. Moreover, complex higher-order systems dynamics and kinetics could be characterized in this manner by essentially applying this approach at a higher scale. Individual components of the system and outputs from these components could be fluorescently labeled and simultaneously observed as with individual parts.

4.2 The digoxin system

The DIG transactivation system is a powerful molecular tool to activate transcription of and quantitatively control a desired output. Furthermore, this system can be designed, as demonstrated in chapter two and three, to target output in a specific cell type. The genetic circuits designs described in

this thesis are proof-of-concept, *i.e.,* the DIG system could be in conjunction with other cell and tissue-specific promoters. Transgenic lines with other cell or tissue specific promoters driving transcription of the DIG dependent transcription factor could be generated to characterization promoters or target output. Moreover, the biomolecule outputted from the DIG system could be changed to explore components of other biological regulatory systems. Other inducible, cell or tissue-specific systems have been utilized to study promoter characteristics, metabolic pathways, or temporal regulation such as circadian rhythms [1][2][3][4]. The DIG transactivation system could feasibly be utilized similarly. However, the circadian rhythm of output needs to be elucidated or addressed before the DIG transactivation system is used in temporal studies.

The separate components and individual parts of the DIG transactivation system should be observed over time to identify what part of the system is temporally regulated. Moreover, through temporal studies of the system possible sources causing oscillatory output over time should be identified. For example, the DIG dependent transcription factor could be tagged with a fluorescence molecule or epitope and the transcription factor level quantified over time. If the DIG dependent transcription factor is temporally regulated, this could extend to the output. In other words, activator regulation is extended to transcriptional regulation of output by means of the promoter the activator binds. Moreover, transgenic lines should be generated with a tagged Gal4-VP16 activator and UAS promoter driving transcription of a

quantifiable output. Observing the UAS/Gal4 system independently could elucidate if temporal regulation is occurring with these components independently and aside from possible temporal regulation of the DIG dependent transcription factor. In the context of the positive feedback system, both transcription factors should be labeled, either with fluorescence or epitope tags, to observe if either transcription factor is temporal regulated or both are. Exploring temporal regulation could go toward inform future approaches. One, the regulation could be accounted for in future designs that incorporate the DIG system. Two, the DIG transactivation system could possibly further tuned or redesigned. If temporal regulation arises from the UAS/Gal-4 components, a possible solution to exclude temporal regulation would be to use a different activator and inducible promoter pair.

4.3 Positive feedback and the digoxin system

The positive feedback genetic circuit was responsive to ligand, had increased sensitivity to input, but lacked amplification of output. Previous mathematically modeling performed by members of the Medford and Prasad labs have shown balancing the promoter strengths of the three promoters within the system is crucial to achieve desired responses of the positive feedback genetic circuit, such as amplification. The ABCG37 promoter was not mathematically characterized, therefore neither was a transfer function for the promoter. However, it is likely this promoter was not balanced within the system given the lack of amplification. The circuit requires a redesign with a different promoter that has been quantified, to achieve a amplified

output. Furthermore, only one of two analyzed transgenic line was responsive. Additional transgenic lines need to be generated and analyzed to explore if the lack of response to ligand and monostablity is a common or rare response of the positive feedback genetic circuit. Furthermore, T2 plants should be generated to observe how amplification, sensitivity, and responsiveness change, if at all, with homozygosity. The availability of several homozygous transgenic lines could be a means to analyze responses of the positive feedback genetic circuit more thoroughly.

The positive feedback genetic circuit output after three days post removal from induction medium is consistent with memory of a transient stimulus. However, decoupling experimental artifact and output is required to elucidate if an overall reduction in output occurs in the memory response, and if the memory persists up to five days. Two approaches could be used to separate signal arising from the memory response and experimental artifact. First, the output of the genetic circuit should be observed at 16 hour intervals. Observing when output peaks could separate the decline in output in the initial 24 hours after removal from induction medium from the oscillations in output observed over time. Second, increasing the number of washes of each sample could remove more residual DMSO from the roots, and thereby minimize or eliminate the increase in GFP signal due to DMSO. Furthermore, different carriers may be tested to see if they could be used instead of DMSO. As a consequence of using these approaches, the output

should principally or solely arise from memory and not as an experimental artifact.

4.4 Modularity and expansion of the digoxin system

The inducible, root-epidermal specific genetic circuit is another tool in the *Arabidopsis* molecular tool kit. The modular nature of the circuit can be leveraged to target output to many tissues and cell-types. Furthermore, the genetic circuit could be expanded upon with regulatory motifs, as demonstrated in chapter three, specifically with the addition of the positive feedback motif to the inducible, root epidermal cell-specific genetic circuit. Other regulatory motifs could be incorporated with the DIG transactivation system to increase the range of possible responses. The DIG transactivation system is a highly valuable resource both in the context of fundamental scientific exploration and designing biosensors with downstream transcriptional regulation with a broad range of responses.

REFERENCES

[1] Radoeva, T., Hove, C.A., Saiga, S., Weijers, D. “Molecular Characterization of Arabidopsis GAL4/UAS Enhancer Trap Lines Identifies Novel Cell Type-Specific Promoters.” *Plant Physiology*, 20 Apr. 2016, doi:10.1104/pp.16.00213.

[2] Hill, C.B., Jha, D.B., Bacic, A.B., Tester, M.B., & Roessner, U.B. Characterization of Ion Contents and Metabolic Responses to Salt Stress of Different Arabidopsis AtHKT1;1 Genotypes and Their Parental Strains.” *Molecular Plant*, vol. 6, no. 2, Mar. 2013, pp. 350–368., doi:10.1093/mp/sss125.

[3] Gendron, J. M., Pruneda-Paz, J.L., Doherty, C.J., Gross, A.M., Kang, S.E., & Kay, S.A. “Arabidopsis Circadian Clock Protein, TOC1, Is a DNA-Binding Transcription Factor.” *Proceedings of the National Academy of Sciences*, vol. 109, no. 8, 21 Feb. 2012, pp. 3167–3172., doi:10.1073/pnas.1200355109.

A.1 Media protocols

A.1.1 Luria-Bertani (LB) Media: Laurel broth for E. coli and A. tumeficians was prepared as follows: To 3800mL of diH20 add 40g Tryptone (Sigma Aldrich, St. Louis, MO), 20g yeast extract (TEKnova, Hollister, CA), and 40g NaCl. pH balance solution to 7 with 3M NaOH. Bring solution up to 4L total, and reverify pH. If solid media is required, add 7.5g of Agar II to 500mL of aliguoted LB media. Autoclave solution.

A.1.2 Murashige and Skoog Basal (MS) Media, solid and liquid: Murashige and Skoog basal media for plants was prepared as follows: To 3200mL of diH20 add 40g sucrose, 17.6g MS basal medium (Sigma Aldrich, St. Louis, MO), and 2g MES (Sigma Aldrich, St. Louis, MO). pH balance media solution to 5.7 with 1M KOH. Bring media solution up to 4L and reverify pH. If solid media is required, aliquot 500mL to bottles and add 3g Plant media Phyto Agar (Plant media a division of bioworld, Dublin, OH) to each aliquot. Autoclave solution.

A.1.3 New Infiltration Media (NIM): New infiltration media for *Agrobacterium*-mediated transformation of Arabidopsis thaliana was prepared as follows: To 3600mL of diH20 add 0.812g of MgCl2, and 200g sucrose. Bring up media solution to 4L. Aliquot solution to 8 bottles each with 500mL of media. Autoclave solution.

A.2 Agrobacterium-Mediated Plant Transformation Protocol

1.) Begin starter culture of GV3101 *Agrobacterium tumefaciens* (NEB, Ipswitch, MA) with desired binary plasmid in LB media (2mL) supplemented with appropriate antibiotic(s) (100mg/ml kanamycin, rif). Allow culture to grow for 2days at 28 ˚C in floor shaker at 220rpm.

2.) Inoculate starter culture in 250mL of LB media supplemented with appropriate antibiotics in 1L flasks. Allow culture to grow overnight at 28˚C in floor shaker at 220 rpm.

3.) Transfer culture to centrifuge bottles, and spin cells at 6000 rpm for 12 minutes at 4°C.

4.) Discard supernatant and resuspend bacterial cells in 500 mL of 'New' infiltration Media containing 25 ﬕL 0.005% Silwet L-77 (Lehle Seeds, Round Rock, TX).

5.) Transfer bacteria-containing infiltration media to 5 L plastic tub.

6.) Inverts pots containing *Arabidopsis thaliana* Columbia ecotype, and dip floral buds into solution for approximately 1 minute. Ensure floral buds are immersed in solution.

7.) After dipping, transfer plants to a flat, with pots laying horizontally.

8.) Cover flat with plastic wrap and move to growth chamber.

9.) After overnight incubation, remove plastic wrap and place plants upright.

A.3 Nucleotide Sequences

Nucleotide Sequences	
Sequence Name	**Nucleotide Sequence (5' to 3')**
UAS Promoter	TTCCTCTCCAAATGAAATGAACTTCCTTATATAGAGGAAGGTCTTGCGCTCGGAGGACAGTACTCCGCTCGGAGGACAGTACTCCGCTCGGAGGACAGTACTCCACTCGGAGGAC AGTACTCCGCTAAGGCATAGTGGGATGTGCGTCATAAGGTTCCAGAGAAATGTTCTGGCACCTGCACTTG
ABCG37 Promoter	CTGAAGTTTATAGAAATTCGTGTATTTTTTTGCCGTTGAAAATGTTTACAGTAGTTAAAAATGAATGATAAACGGTTTTCGTAGATGGTAGAACTCAATCTCTAGACTTGATCCTCT ACACACTCAAACAAACCACTTGGTTACACGTTCTTGGTTGAGCGGATTTCAAATGTGAAACATTTTAACTCATTTTTTTTGGTGAACCTGACCAAAAGAATATTGATTGTACAAGA ACCAAAGTTAAAACTTGTTTTGGTCGTTGTTGAATAAGAAAACCGTGCTCTCAGAAGTCTTTATATAGGCAAACATATAAATTGTGAATTTTGTTTTCTTTATTATCAATTACAATTT ATTTTAAAATATAAATAAAGATATAATAAAGCTATATATTCACGAATAATTCTAAAAAAGTGGTAACAAAAAATGACTTGAAAACATTTTTCTAGTTTCAATTTCAAGTCATTCTAA GAAGTTAAAATTAAGGAATATCACATGCATGAAGTTACATTGAAAGTGTCTTCTTACAAAATAGAAAAAAAAATAAAAAATAAAAACATAGTAAATGTTGTGCATTATTCTTTCAA TAATTAGGGAAAACAATAATACAGTATCTTAGATGTTTATAAGACCAGTATTTATTCATCCAGGCCTAACAAAACTGCCCGCCGCGATTTCTCATGTAAGTATGTAACCTTAGTTAT TAAACAAAATTGCATTTCTTAGACGTTAACCAACTTAATTATTTGTCTTATAGTCAACGTGATTTTTTTACTTTATAATAATGCTTCGATCGTAAGTTGAATCTCAATCAATGTCTTG ATTCCCGAGAGAAGTAGAGCGTTAAAATTGCCGTACAGAACATGATGGAGACATTAATGAATTCAATAAAATCGGCGAAGTTGGAGGTTTAGCTGAGGGAGGCCTGTGTCA TGCTGATTCGTTTAGTTACCAACTCTTCTTTCAAACATCTTTAGTATGTTACTGCTTTCATCCTCTTTATTTGATCTACTGATTATATATATATATTTTTTTTTTACTACATTCGACAT TCGTTGCTAAGAGAATAGAAAACAATTTTCTTTTTCAAAATTATCCTTAGTTTGATATTGGTATGAGAGATTTACATGTAATTTCGACTCTTGAAAATAGATAAATGTATGTTGAGT ATTGTTGACTATAAGAAACTAGATGAATGCATCAAATAGAAGTGTCTTTTTGCTTTATGTACTTCCCCACTTGAAAACATGACGATTCTATGAACGATTTCATTCGTCATAACGATA GAATTAGTTTGTCCTACGAAAACAAATGTAAAACTACTATAATTTTCTTGACGTAATATTGGTGGGTTTCATGCTGCGGCTGTTGGTTGGGAAACGTCCAATATAAACATTTCTAA TTATGTATATACCTGAAAATTTGTAAAAGCTAATCTAAATTGTGAAACATACATAACAAATACTTTAAGAAGAAAAAGTCATCAAGCAAACAAGAAGACATGACCTAGAACCATT TCCGAGACATTATATTAATATAAAATTGATATTCAAATCCTCCACGATCTTACAGCTTTCAACTTTGAGTATTTCAACGTTTTCTCTGATTACACATAAACTAAAATTTGTTTTTTTGT TTTTTTAATAGTTGATAGCATGTGTTAATTATTAGTTTATATCATACAATGACCTTAGAATCCGCAGATCATTTCAAGAACTCAGAACAAAAGCTGATGTATTCACTTACATCATTG AATTGTGTCAGTTTGAGTAACCCATAATATAAATATACGCAAGTCAAAATGCATTGAGCTTATCTCTCATTTAGACTTTAAATTTCTGATTTCAGAGAGGTCAACTCTTTAAGCCAA GCC
2x35S Promoter	AACATGGTGGAGCACGACACTCTCGTCTACTCCAAGAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCTATTGAGACTTTTCAACAAAGGGTAATATCGGGAAACCTCCTCG GATTCCATTGCCCAGCTATCTGTCACTTCATCAAAAGGACAGTAGAAAAGGAAGGTGGCACCTACAAATGCCATCATTGCGATAAAGGAAAGGCTATCGTTCAAGATGCCTCTGC CGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATAACATGGTGGAGCACG ACACTCTCGTCTACTCCAAGAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCTATTGAGACTTTTCAACAAAGGGTAATATCGGGAAACCTCCTCGGATTCCATTGCCCAGCT ATCTGTCACTTCATCAAAAGGACAGTAGAAAAGGAAGGTGGCACCTACAAATGCCATCATTGCGATAAAGGAAAGGCTATCGTTCAAGATGCCTCTGCCGACAGTGGTCCCAAA GATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCC CACTATCCTTCGCAAGACCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGGACACGCTGAAATCACCAGTCTCTCTCTACAAATCTATCTCTCTCGAGCTTTCGCAGATCTGT CGATCGACC
Enhanced GFP	ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATG CCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCC GACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAG TTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTA TATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGG CGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCC GGGATCACTCTCGGCATGGACGAGCTGTACAAG
Matalpha-Gal4-DIG-VP16	ATGAACAAGATCCCAATCAAGGACTTGCTAAATCCACAAATTACTGATGAATTCAAGTCTTCAATATTGGATATAAATAAGAAGCTGTTCTCTATTTGTTGCAACCTCCCAAAGCTT CCGGAATCGTTACTACTGAGGAGGAGGTCGAATTGAAGAGATATATTGGGATTTTTAAGTAGAGCCAACAAAAACAGGAAGATTATGAAGTTGTTAAGTTCTATTGAACAGGCT TGTGATATTTGTAGATTGAAGAAGCTCAAGTGTAGTAAAGAGAAACCTAAGTGCGCTAAGTGTCTTAAAAATAACTGGGAGTGCAGATATTCTCCTAAGACTAAGAGATCACCTC TTACTAGGGCTCATCTCACAGAGGTGGAGTCTAGGCTTGAGAGATTGGAGCAGTTGTTCCTTTTGATTTTTCCAAGAGAAGATCTTGATATGATTCTTAAGATGGATTCTCTTCAA GATATTAAGGCTCTTCTTGGAGGTTCTGGCGGATCCATGAATGCTAAAGAAATTGTTGTCCACGCTCTCCGTCTGCTGGAAAATGGCGATGCCCGTGGTTGGTCTGACCTGTTTCA CCCGGAAGGCGTGCTGGAATATCCGTACCCACCGCCGGGCTATAAAACCCGTTTCGAAGGTCGCGAAACGATTTGGGCGCACATGCGTCTGTTCCCGGAATATATGACCATCCGC TTTACGGATGTCCAGTTCTACGAAACCGCCGATCCGGACCTGGCAATCGGCGTATTTCATGGTGACGGTGTTCTCACCGCAAGCGGCGGTAAACTGGCGTACGATTATATTGCCG TTTGGCGTACGCGCGACGGTCAGATTCTGCTGTACCGTCTCTTTTTCAACCCGCTGCGTGTCCTGGAACCACTGGGCGGTGTGGAAGCAGCTGCGAAAATTGTTCAAGGCGCGTC TGCAGATGGTAGTGCGTACAGCCGCGCGCGTACGAAAAACAATTACGGGTCTACCATCGAGGGCCTGCTCGATCTCCCGGACGACGACGCCCCCGAAGAGGCGGGGCTGGCGG CTCCGCGCCTGTCCTTTCTCCCCGCGGGACACACGCGCAGACTGTCGACGGCCCCCCCGACCGATGTCAGCCTGGGGGACGAGCTCCACTTAGACGGCGAGGACGTGGCGATGG CGCATGCCGACGCGCTAGACGATTTCGATCTGGACATGTTGGGGGACGGGGATTCCCCGGGTCCGGGATTTACCCCCCACGACTCCGCCCCCTACGGCGCTCTGGATATGGCCGA CTTCGAGTTTGAGCAGATGTTTACCGATGCCCTTGGAATTGACGAGTACGGTGGGTAA
Gal4-VP 64 Transcription factor	ATGAAGTTGTTAAGTTCTATTGAACAGGCTTGTGATATTTGTAGATTGAAGAAGCTCAAGTGTAGTAAAGAGAAACCTAAGTGCGCTAAGTGTCTTAAAAATAACTGGGAGTGC AGATATTCTCCTAAGACTAAGAGATCACCTCTTACTAGGGCTCATCTCACAGAGGTGGAGTCTAGGCTTGAGAGATTGGAGCAGTTGTTCCTTTTGATTTTTCCAAGAGAAGATCT TGATATGATTCTTAAGATGGATTCTCTTCAAGATATTAAGGCTCTTCTTACTGGTTTATTCGTTCAAGATAATGTGAATAAGGATGCAGTGACTGATAGACTTGCATCAGTTGAAA CTGATATGCCTCTTACATTAAGACAACACAGGATTTCTGCTACTTCAAGTTCTGAAGAAAGTTCTAATAAGGGTCAAAGGCAATTAACCGTTTCTGCCGACGCGCTGGACGATTTC GATCTCGACATGCTGGGTTCTGATGCCCTCGATGACTTTGACCTGGATATGTTGGGAAGCGACGCATTGGATGACTTTGATCTGGACATGCTCGGCTCCGATGCTCTGGACGATTT CGATCTCGATATGTTAGGTTCCCCTAAGAAGAAAAGGAAGGTGGGT
TNOS Terminator	GATCGTTCAAACATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAA TGCATGACGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGG TGTCATCTATGTTACTAGATC
tADH Terminator	AGCCATTCTCTCGCAGATGATGTTCACTTTGTGTTTTACTTCCTTTATGCATTCACAGCAATAAAAGAAAGAAATCTCCATCGCTTTTGGTTTTCTTCTCTGTCTTAAGTTAGTCGTTT TCGTGTCTAATCTATTACTTATCATTGTAATAGACTCTTCTTCTATTGAGATTTGAATATAAACTAAAACACATTCCATTTTACTGTGTTCTCAACATTCAGAATGCAAACGGACTAA CCGTAGTGACTCG
35S Terminator	GCGGGACTCTGGGGTTCGGATCGATCCTCTAGCTAGAGTCGATCGACAAGCTCGAGTTTCTCCATAATAATGTGTG
Transcription Block	AATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCT GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCT
Kanamycin Resistance (aadA)	ATGGCTAAAATGAGAATATCACCGGAATTGAAAAAACTGATCGAAAAATACCGCTGCGTAAAAGATACGGAAGGAATGTCTCCTGCTAAGGTATATAAGCTGGTGGGAGAAAA TGAAAACCTATATTTAAAAATGACGGACAGCCGGTATAAAGGGACCACCTATGATGTGGAACGGGAAAAGGACATGATGCTATGGCTGGAAGGAAAGCTGCCTGTTCCAAAGG TCCTGCACTTTGAACGGCATGATGGCTGGAGCAATCTGCTCATGAGTGAGGCCGATGGCGTCCTTTGCTCGGAAGAGTATGAAGATGAACAAGCCCTGAAAAGATTATCGAGC TGTATGCGGAGTGCATCAGGCTCTTTCACTCCATCGACATATCGGATTGTCCCTATACGAATAGCTTAGACAGCCGCTTAGCCGAATTGGATTACTTACTGAATAACGATCTGGCC GATGTGGATTGCGAAAACTGGGAAGAAGACACTCCATTTAAAGATCCGCGCGAGCTGTATGATTTTTTAAAGACGGAAAAGCCCGAAGAGGAACTTGTCTTTTCCCACGGCGAC CTGGGAGACAGCAACATCTTTGTGAAAGATGGCAAAGTAAGTGGCTTTATTGATCTTGGGAGAAGCGGCAGGGCGGACAAGTGGTATGACATTGCCTTCTGCGTCCGGTCGATC AGGGAGGATATCGGGGAAGAACAGTATGTCGAGCTATTTTTTGACTTACTGGGGATCAAGCCTGATTGGGAGAAAATAAAATATTATATTTTACTGGATGAATTGTTT
Kanamycin Resistance (nptII)	ATGGGGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTC CGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTCCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTT CCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGTAT CCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCT TGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCACATGGCG ATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATAT TGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTC

A.4 Primer Sequences

Primer Number	Primer Sequence (5' to 3')	Primer Number	Primer Sequence (5' to 3')
1	GATTACGAATTCGAGCTCTTCCT CTCCAAATG	18	CGCTCAAccAAGAACGTGTAACC AAG
2	CATTGTTTGACGTCCAAGTGCAG GTG	19	GAA CGT CAG AAG CCG ACT GCA CTA TA
3	GCTCTTTTCTCTTAGGTTTACCCG CCAATA	20	GAA CCC TGT GGT TGG CAT GC
4	GTTTCACATTTGAAATCCGCTCA ACCAAGAAC	21	GTC CAG GAG CGC ACC ATC TTC
5	GCTTCATGATCAAAAATGAACAA GATCCCAATCAAG	22	CGG ACT AAC CGT AGT GAC TCG
6	ACTTGACTTTAAACTAGATTACCC ACCGTACTGG	23	CTT GCC GGT GGT GCA GAT GAA CTT
7	ATCCTCCTGGGTACCCTGAAGTT TATAGAAATTCGTGT	24	CGCTCAAccAAGAACGTGTAACC AAG
8	TAGGTGGTTAATTAAGGCTTGGC TTAAAGAGTTGACCTCTC	25	TTG TAC AGC TCG TCC ATG CCG
9	GAA CCC TGT GGT TGG CAT GC	26	CAGCTAAACCTCCAACTTCGCCG
10	GTC CAG GAG CGC ACC ATC TTC	27	GTCCGCTCTACCGAAAGTTACGG
11	CGG ACT AAC CGT AGT GAC TCG	28	GT AAC ATA GAT GAC ACC GCG CGC
12	CCG CCG CGA TTT CTC ATG TAA G	29	CTC TGC GAT CGC TTC CTC TCC
13	CAT GCT GCG GCT GTT GGT TG	30	GGATGCAGTGACTGATAGACTT GC
14	GAG AGG TCA ACT CTT TAA GCC AAG C	31	CTC CGA TGC TCT GGA CGA TTT CG
15	CGT GGT TGG TCT GAC CTG TTT C	32	GGCAATTAACCGTTTCTGCCGAC
16	CGC GCT AGA CGA TTT CGA TCT G	33	GAG GAC AGT ACT CCG CTA AGG C
17	CTT GCC GGT GGT GCA GAT GAA CTT	34	CGCTCAACCAAGAACGTGTAAC CAAG

A.5 MATLAB Code

```
Copyright Sara Oehmke 2020
FolderTS = '';
if ~isdir(FolderTS)
  errorMessage = sprintf('Error: The following folder does not exist:\n%s', FolderTS);
  uiwait(warndlg(errorMessage));
  return;
end
filePatternTS = fullfile(FolderTS, '*.tiff');
theFilesTS = dir(filePatternTS);
for h = 1 : length(theFilesTS)
   NamesTS = {theFilesTS.name};
end
  TotalFileNamesTS= strcat(FolderTS,NamesTS);
for g = 1:length(TotalFileNamesTS)
  AllimagesTS{g} = imread(TotalFileNamesTS{g});
end
AllimagesTS2 = cellfun(@im2double,AllimagesTS, 'uniform', 0);
AllimagesTS3 = cell2mat(AllimagesTS2);
FolderBGS= '';
if ~isdir(FolderBGS)
  errorMessage = sprintf('Error: The following folder does not exist:\n%s', FolderBGS);
  uiwait(warndlg(errorMessage));
  return;
end
filePatternBGS = fullfile(FolderBGS,'*.tiff');
theFilesBGS = dir(filePatternBGS);
for k = 1 : length(theFilesBGS)
   NamesBGS = {theFilesBGS.name};
end
  TotalFileNamesBGS= strcat(FolderBGS,NamesBGS);
for n = 1:length(TotalFileNamesBGS)
  AllimagesBGS{n} = imread(TotalFileNamesBGS{n});
end
AllimagesBGS1 = cellfun(@im2double,AllimagesBGS, 'uniform', 0);
AllimagesBGS2 = cell2mat(AllimagesBGS1);
for m = 1 : length(AllimagesYFP5)
  T2{m} = corrcoef(AllimagesYFP5{m},AllimagesGFP5{m});
end
R = corr2(AllimagesBGS2, AllimagesTS3)
equation= generatinggraph(AllimagesBGS2,AllimagesTS3)
Linearcoeff=equation(1,1)
Intercept=equation(1,2)
FolderTST = '';
if ~isdir(FolderTST)
  errorMessage = sprintf('Error: The following folder does not exist:\n%s', FolderTST);
  uiwait(warndlg(errorMessage));
  return;
end
filePatternTST = fullfile(FolderTST, '*.tif'); % Change to whatever pattern you need.
theFilesTST = dir(filePatternTST);
for k = 1 : length(theFilesTST)
  baseFileNameTST = theFilesTST(k).name;
  fullFileNameTST= fullfile(FolderTST, baseFileNameTST);
```

```
  fprintf(1, 'Now reading %s\n', fullFileNameTST);
end
for k = 1 : length(theFilesTST)
   NamesTST = {theFilesTST.name};
   FolderTST = {theFilesTST.folder};
end
  TotalFileNamesTST= strcat(FolderTST,NamesTST)
for b = 1:length(TotalFileNamesTST)
  AllimagesTST{b} = imread(TotalFileNamesTST{b});
end
AllimagesTST = cellfun(@im2double,AllimagesTST, 'uniform', 0)
FolderBGST = '';
if ~isdir(FolderBGST)
  errorMessage = sprintf('Error: The following folder does not exist:\n%s', FolderBGST);
  uiwait(warndlg(errorMessage));
  return;
end
filePatternBGST = fullfile(FolderBGST, '*.tif'); % Change to whatever pattern you need.
theFilesBGST = dir(filePatternBGST);
for k = 1 : length(theFilesBGST)
  baseFileNameBGST = theFilesBGST(k).name;
  fullFileNameBGST = fullfile(FolderBGST, baseFileNameBGST);
  fprintf(1, 'Now reading %s\n', fullFileNameBGST);
end
for k = 1 : length(theFilesBGST)
   NamesBGST = {theFilesBGST.name};
   FolderBGST = {theFilesBGST.folder};
end
  TotalFileNamesBGST= strcat(FolderBGST, NamesBGST)
for b = 1:length(TotalFileNamesBGST)
  AllimagesBGST{b} = imread(TotalFileNamesBGST{b});
end
AllimagesBGST = cellfun(@im2double,AllimagesBGST, 'uniform', 0)
for i = 1:length(AllimagesBGST)
  AllimagesBGST{1, i} =  AllimagesBGST{1,i}-Intercept;
  AllimagesBGST{1, i} =  AllimagesBGST{1,i}*Linearcoeff;
end
FluorescenceImage = cellfun(@minus,AllimagesTST,AllimagesBGST,'Un',0)
Meannomask = AFremoval2(FluorescenceImage);
MeanNOLOG = AFremoval3(FluorescenceImage);
MeanLOG = AFremoval4(FluorescenceImage);
TotalMean2 = struct('mean',[Meannomask], 'meanMask', [MeanNOLOG], 'meanLog',
[MeanLOG]);
TotalMean = structfun(@(fld) fld(:), TotalMean2,'UniformOutput', false);
TotalMean3 = struct2table(TotalMean);
filename = '.xlsx';
writetable(TotalMean3,filename,'Sheet',1)

%Functions
function equation =createfigure(AllimagesBGS1, AllimagesTS3)
 figure1 = figure;
 axes1 = axes('Parent',figure1,'Position',[0.13 0.5675 0.775 0.3575]);
hold(axes1,'on');
 scatter1 = scatter(AllimagesBGS1(:), AllimagesTS3(:),'DisplayName','data1','Parent',axes1);
 xdata1 = get(scatter1, 'xdata');
```

```
ydata1 = get(scatter1, 'ydata');
xdata1 = xdata1(:);
ydata1 = ydata1(:);
 nanMask1 = isnan(xdata1(:)) | isnan(ydata1(:));
if any(nanMask1)
    warning('IgnoringNaNs', ...
        'Data points with NaN coordinates will be ignored.');
    xdata1(nanMask1) = [];
    ydata1(nanMask1) = [];
end

axesLimits1 = xlim(axes1);
xplot1 = linspace(axesLimits1(1), axesLimits1(2));
 set(axes1,'position',[0.1300    0.5811    0.7750    0.3439]);
residAxes1 = axes('position', [0.1300    0.1100    0.7750    0.3439], ...
    'parent', gcf);
savedResids1 = zeros(length(xdata1), 1);
[sortedXdata1, xInd1] = sort(xdata1);
 coeffs1 = cell(1,1);
 fitResults1 = polyfit(xdata1,ydata1,1);
equation=fitResults1;
yplot1 = polyval(fitResults1,xplot1);
 fittypesArray1(1) = 2;
 Yfit1 = polyval(fitResults1,xdata1);
resid1 = ydata1 - Yfit1(:);
savedResids1(:,1) = resid1(xInd1);
savedNormResids1(1) = norm(resid1);
 coeffs1{1} = fitResults1;
 fitLine1 = plot(xplot1,yplot1,'DisplayName','   linear','Tag','linear',...
    'Parent',axes1,...
    'Color',[0.929 0.694 0.125]);
 setLineOrder(axes1,fitLine1,scatter1);
 residPlot1 = plot(sortedXdata1,savedResids1,'.','parent', residAxes1);
set(residPlot1(1), 'color', [0.929 0.694 0.125]);
title(residAxes1, 'residuals');
 showNormOfResiduals(residAxes1,fittypesArray1,savedNormResids1);
 showEquations(fittypesArray1,coeffs1,2,axes1);
 legend(axes1,'show');
 function setLineOrder(axesh1, newLine1, associatedLine1)
hChildren = get(axesh1,'Children');
hChildren(hChildren==newLine1) = [];
lineIndex = find(hChildren==associatedLine1);
hNewChildren = [hChildren(1:lineIndex-1);newLine1;hChildren(lineIndex:end)];
set(axesh1,'Children',hNewChildren);
function showNormOfResiduals(residaxes1, fittypes1, normResids1)
txt = cell(length(fittypes1) ,1);
for i = 1:length(fittypes1)
    txt{i,:} = getResidString(fittypes1(i),normResids1(i));
end
axesunits = get(residaxes1,'units');
set(residaxes1,'units','normalized');
text(.05,.95,txt,'parent',residaxes1, ...
    'verticalalignment','top','units','normalized');
set(residaxes1,'units',axesunits);
 function [s1] = getResidString(fittype1, normResid1)
```

```
switch fittype1
   case 0
      s1 = getString(message('MATLAB:graph2d:bfit:ResidualDisplaySplineNorm'));
   case 1
      s1 = getString(message('MATLAB:graph2d:bfit:ResidualDisplayShapepreservingNorm'));
   case 2
      s1 = getString(message('MATLAB:graph2d:bfit:ResidualDisplayLinearNorm',
num2str(normResid1)));
   case 3
      s1 = getString(message('MATLAB:graph2d:bfit:ResidualDisplayQuadraticNorm',
num2str(normResid1)));
   case 4
      s1 = getString(message('MATLAB:graph2d:bfit:ResidualDisplayCubicNorm',
num2str(normResid1)));
   otherwise
      s1 = getString(message('MATLAB:graph2d:bfit:ResidualDisplayNthDegreeNorm',
fittype1-1, num2str(normResid1)));
end
 n = length(fittypes1);
txt = cell(length(n + 1) ,1);
txt{1,:} = ' ';
for i = 1:n
   txt{i + 1,:} = getEquationString(fittypes1(i),coeffs1{i},digits1,axesh1);
end
text(.05,.95,txt,'parent',axesh1, ...
   'verticalalignment','top','units','normalized');
 if isequal(fittype1, 0)
   s1 = 'Cubic spline interpolant';
elseif isequal(fittype1, 1)
   s1 = 'Shape-preserving interpolant';
else
   op = '+-';
   format1 = ['%s %0.',num2str(digits1),'g*x^{%s} %s'];
   format2 = ['%s %0.',num2str(digits1),'g'];
   xl = get(axesh1, 'xlim');
   fit =  fittype1 - 1;
   s1 = sprintf('y =');
   th = text(xl*[.95;.05],1,s1,'parent',axesh1, 'vis','off');
   if abs(coeffs1(1) < 0)
      s1 = [s1 ' -'];
   end
   for i = 1:fit
      sl = length(s1);
      if ~isequal(coeffs1(i),0) % if exactly zero, skip it
         s1 = sprintf(format1,s1,abs(coeffs1(i)),num2str(fit+1-i), op((coeffs1(i+1)<0)+1));
      end
      if (i==fit) && ~isequal(coeffs1(i),0)
         s1(end-5:end-2) = []; % change x^1 to x.
      end
      set(th,'string',s1);
      et = get(th,'extent');
      if et(1)+et(3) > xl(2)
         s1 = [s1(1:sl) sprintf('\n     ') s1(sl+1:end)];
      end
   end
```

```
    if ~isequal(coeffs1(fit+1),0)
        sl = length(s1);
        s1 = sprintf(format2,s1,abs(coeffs1(fit+1)));
        set(th,'string',s1);
        et = get(th,'extent');
        if et(1)+et(3) > xl(2)
            s1 = [s1(1:sl) sprintf('\n     ') s1(sl+1:end)];
        end
    end
    delete(th);
    % Delete last "+"
    if isequal(s1(end),'+')
        s1(end-1:end) = []; % There is always a space before the +.
    end
    if length(s1) == 3
        s1 = sprintf(format2,s1,0);
    end

end

function [meanIntSampleNOLOG] = AFremoval3(GFPunstructured)
 for l = 1 : length(GFPunstructured)
[binarylevelGFPNOLOG, EM] = graythresh(GFPunstructured{l});
BW = imbinarize(GFPunstructured{l});
BW2 = imfill(BW,'holes');
maskedImageGFP2 = (GFPunstructured{l}); % Initialize
maskedImageGFP2(~BW) = 0; % Erase everything outside the mask.
meanIntSampleNOLOG(l)= mean2(maskedImageGFP2);
end

function [meanIntSampleNomask, meanIntSampleNOLOG, meanIntSampleLOG] =
AFremoval(GFPunstructured)
 for k = 1 : length(GFPunstructured)
meanIntSampleNomask(k) = mean2(GFPunstructured{k});
end

function [meanIntSampleLOG] = AFremoval4(FluorescenceImage)
 for M = 1 : length(FluorescenceImage)
GFPlog= 10*log(1+ FluorescenceImage{M});
[binarylevelGFP, EM] = graythresh(GFPlog)
BW3 = imbinarize(GFPlog);
BW4 = imfill(BW3,'holes');
maskedImage = GFPlog; % Initialize
maskedImage(~BW3) = 0; % Erase everything outside the mask.
meanIntSampleLOG(M) = mean2(maskedImage);
end
```