More than Meets the Eye: Emergent Properties of Transcription Factors Networks in Arabidopsis

Durreshahwar Muhammad^{*a}, Selene Schmittling^{*b}, Cranos Williams^b, Terri A. Long^{1a}

^a Plant and Microbial Biology, North Carolina State University, Raleigh, NC, USA

^b Electrical and Computer Engineering, North Carolina State University, Raleigh, NC, USA

*These authors contributed equally to this work

¹Corresponding author

ABSTRACT

Uncovering and mathematically modeling Transcription Factors Networks (TFNs) are the first steps in engineering plants with traits that are better equipped to respond to changing environments. Although several plant TFNs are well known, the framework for systematically modeling complex characteristics such as switch-like behavior, oscillations, and homeostasis that emerge from TFNs remain elusive. This review highlights literature that provides, in part, experimental and computational techniques for characterizing TFNs. This review also outlines methodologies that have been used to mathematically model the dynamic characteristics of TFNs. We present several examples of TFNs in plants that are involved in developmental and stress response. In several cases, advanced algorithms capture or quantify emergent properties that serve as the basis for robustness and adaptability in plant responses. Increasing the use of mathematical approaches will shed new light on these regulatory properties that control plant growth and development, leading to mathematical models that predict plant behavior.

Keywords: Transcription factor networks (TFNs), emergent properties, mathematical modeling

1. INTRODUCTION

Plants are sessile organisms that utilize robust regulatory networks, which enable them to thrive in changing environments. Understanding how plants respond to different stresses and constraints is critical for engineering crops that can withstand changing conditions. Transcription factors, which are key proteins responsible for the regulation of gene expression, are often controlled by internal and external biochemical signals. This response initiates a cascade of regulatory processes that allows the plant to respond and adapt effectively. The interaction

between transcription factors and the proteins they regulate lead to emergent properties. Emergent properties refer to the behavior of complex systems that cannot be predicted based on the behavior of its individual components [1]. Examples of emergent behavior in TFNs include bistability, oscillatory or switch-like changes in transcript concentrations. To this end, understanding transcriptional regulation and the emergent properties that can arise from this regulation is a first and necessary step to unraveling the complex interactions at all levels of regulation (e.g., post-transcriptional and translational modifications) [2]. *Arabidopsis thaliana* serves as an excellent model organism to examine the transcriptional landscape of plants because its developmental schemes are well studied. Its genome is also relatively small and fully sequenced. *Arabidopsis* roots in particular have a simple developmental structure that is regulated by several known transcription factor networks. Ongoing research has shown that transcription factors and their associated regulatory networks contain network motifs that may be responsible for the emergent properties seen in gene expression, phenotypic attributes, and overall system behavior [3,4]. A network motif is a specific layout of transcription factors (nodes) and interactions (edges) of a small subset of nodes and edges for a much larger network. These motifs correlate with specific forms of behavior, i.e., emergent properties, in the biological and mathematical systems in which they are found. These motifs have been associated with different types of emergent behavior including bistability, homeostasis, and sustained oscillations [5,6]. In practice, transcription factor networks rarely contain only one motif. Instead, they tend to be made up of motifs embedded within and linked to each other. The complex nature of the interactions that result from these connections is difficult to predict. The link between these motifs, the biological interactions they represent and the resulting phenotypic behavior is often difficult to quantify and confirm.

Efforts to mathematically model such motifs have begun to shed new light on the dynamics of TFNs and their resulting molecular and phenotypic outcomes. This is because network motif models can be utilized in broader mathematical contexts to describe larger more complex regulatory systems. Ideally, once a modeling paradigm has been generated that can reasonably explain the interactions of key regulators and their targets, *in silico* data generated by the model can be verified against known information about the system under investigation or by using additional experimental data. The benefit of such models is that they can be used to predict plant behaviors and to engineer plants with more robust responses to stress, larger yield and other desirable traits.

In this review, we discuss emergent properties as they relate to gene expression in plant systems, computational and experimental approaches that have been used to build transcription factor networks (TFNs), network motifs that have been shown in the literature to lead to these types of properties, and systems

approaches that use dynamical models of these motifs to verify, *in silico*, their existence. We then provide examples and discuss known TFNs in Arabidopsis, particularly those involved in development and cell maintenance, abiotic stress response, and nutritional stress. It is hypothesized that these TFNs exhibit emergent properties and/or contain network motifs that contribute to the complex characteristics seen in the regulatory response.

While a number of network motifs have been identified, mathematical modeling of such transcription networks is particularly challenging because it requires expertise across diverse disciplines. The work outlined in this review provides an interdisciplinary roadmap for better understanding the characteristics of TFNs, where combinations of these regulatory motifs lead to functionality that typically extends beyond the sum of the regulatory components. Further efforts to merge expertise in specific biological processes as well as computational acumen will lead to the development of new interdisciplinary tools and technological advances in the area.

2. EMERGENT PROPERTIES ASSOCIATED WITH NETWORK MOTIFS AND THEIR BIOLOGICAL REPRESENTATIONS

Emergent properties arise in nonlinear dynamical systems. Biological systems are highly dynamic and nonlinear and therefore frequently give rise to emergent properties. They can be seen at all scales in the biological world from ecosystems that include fish schools and the flying formation of birds [7] to the emergence of bistability in transcripts responsible for stem cell maintenance [8]. Transcription factor networks and the relationships they describe are excellent examples of dynamic nonlinear systems in biology. These networks can be represented as undirected (weighted and unweighted) or directed graphs [9-11]. These graphs include nodes, which represent transcription factors and their targets, and edges that represent a relationship between specific nodes [9]. Relationships between nodes are indicated with symbols for activation (a pointed arrow) and inhibition (a blunt arrow). Bistability, oscillations and switch-like behavior can only be identified in systems where the relationships (inhibition or activation) between transcription factors and other genes are known or assumed to be known since these properties arise from the combination of activation and inhibition that occur in the system.

Graphical models of relationships between TFs and their targets are often described in terms of network motifs (Figure 1). An analysis of network motifs that spanned disciplines [12,13] discovered 13 three-component network motifs that are statistically overrepresented across all types of networks, including biological networks, compared to motifs found in random networks. In transcriptional control, a limited set of network motifs have been described [5]

(Figure 2). While most of these motifs span sensory response and developmental processes, some are almost exclusively associated with developmental processes [5]. Common motifs found in transcriptional regulation include simple regulation, positive and negative autoregulation, positive and negative feedback loops, and feed-forward loops (Figure 2).

Transcriptional feedback loops are network motifs in which a gene product in the transcriptional cascade goes back and regulates the activation or inhibition of another gene in the cascade. Negative feedback results in an inhibition of the gene while positive feedback results in the activation of that gene. The presence of multiple negative feedback loops can impact system behavior. Negative feedback loops occur when the number of inhibitory edges in the feedback structure is odd. If the number of inhibitory edges is even, this results in a positive feedback loop. Feedback can lead to bifurcations, which are drastic changes in system characteristics (e.g. concentration dynamics, changes in stability, and emergence of steady states) when small changes in the system parameters occur [14]. Bifurcations can lead to emergent properties such as bistability (multiple steady states) or sustained oscillations [14,15]. Switch-like behavior in biological systems has been associated with biological processes such as cell differentiation [16,17]. Negative feedback loops and nested feedback loops have been associated with sustained oscillations and periodic behavior over time [6,18]. Gene oscillations are also a key component in circadian processes [18].

Autoregulation is a subclass of feedback loops in which the product of a gene regulates its own transcription. Many systems incorporate positive or negative autoregulation. Positive autoregulation is associated with bistability, where the transcript can attain one of two stable states. Negative autoregulation has been associated with three types of behaviors: (1) homeostasis, where transcript levels are kept within a tight range; (2) sustained oscillations, where transcript levels cycle through a range of concentrations, exhibiting periodic behavior over time and (3) robustness to noise, where transcript levels exhibit little change in the presence of significant internal and external perturbations [3]. Negative autoregulation has also been shown to produce faster transcriptional response [5]. Two large transcription factor families involved in the response of *Arabidopsis* to various biotic and abiotic stressors, NAC and WRKY, are well known for their autoregulatory behavior [19,20]. It is likely that autoregulation plays a key role in the ability of transcription factors to respond to multiple stresses.

Feed-forward loops are a common motif associated with specific types of emergent properties. Feed forward loops are ones in which a transcription factor regulates another and they both regulate the production of a third (Figure 2). Feed-forward loops involving three nodes were characterized in Mangan, 2003.

The presence of three nodes and two possible interactions (inhibition or activation) leads to eight possible types of feed-forward loops, two of which occur much more frequently than the others [5]. Feed-forward loops are categorized as coherent or incoherent based on the consistency of signals feeding into the output node (Figure 3a). Feed-forward loops control the response of the tertiary (output) node in a variety of ways. The output can be delayed upon stimulation of the primary node or when stimulation stops. The output may also depend on the logic associated with the input of the primary and secondary nodes (Figure 3b). This may lead to pulse-like behavior for the output of the tertiary node or response acceleration [5].

In cases where multiple transcription factors regulate a target gene, the output of the system will depend on the Boolean logic of the inputs (Figure 3b). Relationships seen in a transcription factor network can be described based on simple Boolean logic rules, which assume that the concentration of each node is present ('1') or not present ('0'). For the simple case where two TFs regulate a single target, these rules include logical AND (both regulators needed), logical OR (either regulator needed), and logical Exclusive OR (XOR) (either regulator, but not both).

Transcription factor networks are composed of interconnected network motifs (Figure 1). Determining the behavior of TFN based solely on individual motifs presents a challenge. Generating a mathematical model of the system that includes each motif within the context of the network enables more quantitative investigation of the existence of complex behaviors such as bistability, sustained oscillations, or bifurcation. These models can be used to uncover or confirm components and interactions that lead to complex behavior in a quantitative manner by comparing it to experimental data. The resulting quantitative knowledge gives researchers the ability to predict plant behaviors or engineer plant systems that can respond to stresses in novel ways. In the next sections we discuss tools for generating transcription factor networks that can be used to explore emergent properties, modeling techniques for identifying these properties, and review recent examples of TFNs that exhibit behavior or contain network motifs associated with bistability, oscillations or switch-like behavior that have not yet been modeled.

3. IDENTIFYING TRANSCRIPTION FACTOR NETWORKS AND THEIR EMERGENT PROPERTIES

Characterizing the nature of transcription factor networks starts with the elucidation of specific nodes involved in the transcriptional process of interest. These nodes include the transcription factors and the genes they regulate. In addition, the interactions between nodes, also known as activation or inhibition

edges, must be identified (or in some cases surmised). Once a comprehensive network has been determined, the emergent behaviors can be examined by quantifying the activity using mathematical models.

3.1 FINDING NODES: Transcription Factors and their Targets

Transcription factors and their associated targets can be identified via literature curation, experimental techniques, or computational approaches. Experimental approaches used to identify transcription factor regulators and/or their direct targets include Yeast One-Hybrid (Y1H) analysis, Protein Binding Microarrays (PBM), DHS mapping with genomic footprinting, and chromatin immunoprecipitation sequencing (ChIP-seq). Computational approaches for identifying transcription factors and their targets include cis- and trans- element analysis [21,22] and temporal analysis of high throughput transcriptional data[23,24].

3.1.1 Experimental Approaches for Identifying TFs and their Targets

Transcription factors control the expression of their targets by directly binding to the promoter region of target genes. Y1H analysis is an experimental tool used to identify protein-DNA interactions, where *in vivo* interactions between the transcription factor (prey) and cis-acting element sequences (bait) are detected in yeast reporter strains[25]. In general, however, protein-DNA binding is not limited to transcription factors. To restrict binding results to transcription factors, several labs have generated libraries that only contain transcription factors as bait [26-28]. As with most *in vitro* experiments, the biological relevance of the interactions identified using this approach must be verified *in vivo*. In addition, Y1H analysis cannot detect TF-DNA interactions that require TF binding partners or complex formation. Despite these caveats, Y1H analysis has been very helpful for identifying many potential TF targets *in vitro* [26,27]. These targets can subsequently be used as nodes in the TF network.

Similar to Y1H analysis, protein-binding microarrays (PBMs) can also be used to identify protein-DNA interactions. This approach utilizes TFs tagged with an epitope that is applied to a microarray for binding to targets. After binding, the microarray is washed and probed against the primary antibody specific to the tag [29]. Binding of a TF to the PBM, however, does not conclusively signify that the binding occurs *in planta*. Another technique that can be used to identify sequences to which proteins bind is DNase hypersensitivity sites mapping. DNase hypersensitivity sites represent areas in the DNA that have a high probability of cleavage by DNase I. Cleaved sites tend to occur close to the location where proteins bind to the DNA. Thus, DNase I hypersensitivity mapping (DHS) coupled with genetic footprinting can be used to identify cis-regulatory

elements and network connectivity [13,30,31]. Chromatin immunoprecipitation followed by microarray hybridization (ChIP-chip) or sequencing (ChIP-SEQ) is a common experimental approach for identifying direct targets of transcription factors *in vivo*. This method focuses solely on a single transcription factor and its targets. ChIP-Chip/SEQ uses transgenic plants expressing the transcription factor of interest to determine protein-DNA interactions [32]. DNA acquired from chromatin immunoprecipitation is often low quality and quantity, presenting a problem when hybridized or sequenced. After identifying direct targets using Y1H, DHS and genetic footprinting, PBM or CHIP-Chip/SEQ, results are often further analyzed using gene ontology analysis. This helps to determine biological functions associated with groups of TF targets. Comparisons of TFs targets specifically to transcription factor databases (DATD; <http://datf.cbi.pku.edu.cn/>, or AtTFDB; <http://arabidopsis.med.ohio-state.edu/AtTFDB/>) can confirm TF-TF relationships. This allows for the start of a linear TFN with two levels. Adding additional nodes to the network requires literature searches and an iteration of protein-DNA binding assays on the newly identified TFs.

3.1.2 Computational Approaches for Predicting TFs and Their Targets

Computational tools which identify plausible nodes in a network include searches of the promoter space of single genes or groups of co-expressed genes for statistically significant conserved cis-regulatory elements that bind specific TFs [33-35]. Several databases are available for this task including AGRIS, AthaMAP, AtProbe, and DATF. Programs such as the Athena analysis suite [34], Promomer [35], and MEME [36] find statistically overrepresented cis-regulatory motifs in a group of related genes, or in some cases, the entire genome. They return known transcription binding sites and/or over-represented sequences of a given length. This information is then matched to specific transcription factors, revealing the regulators of specific targets. Several other open source software programs, such as MEME, that can help identify novel or known promoter elements from a given set of genes [36].

High-throughput transcriptomic data can be used to predict possible transcription factors and their targets computationally. Time-lagged correlation analysis has been used in combination with time-course transcriptomic data to identify potential genes that activate target genes [23]. Koryachko, et al., developed the Cluster and Differential Alignment Algorithm (CDAA), which used a combinatorial computational approach to build a network of genes involved in iron deficiency response from time course microarray data. While these experimental and computational approaches can identify potential gene regulators and their targets, further analysis is needed to determine whether those regulators activate or inhibit the expression of the genes they regulate.

3.2. FINDING EDGES - Activation and Inhibition

Determining the type of interaction (activation or inhibition) between nodes in a TFN is a necessary step in formulating models that capture complex behavior. A combination of known interactions from literature, experimental techniques, and modeling approaches can be used to define edges between genes. Here we discuss existing experimental and computational methods for determining the type of influence between transcription factors and their targets.

3.2.1 Computational Analysis of Experimental Data

Experimental data used to determine the nature of regulation often involve transcript analysis such as qRT-PCR, microarray analysis, and RNA-Seq of T-DNA insertion lines, RNAi lines, and overexpression transgenic lines. However, targets may have several TFs regulating their activity in similar or opposing manners. Li et al. found a negative correlation between the number of TFs that bind a single target and the effect of single regulators on target transcript abundance [37]. Thus, it is often necessary to study higher order mutants of multiple regulator TFs to understand how they regulate target gene expression in tandem.

Computational approaches that predict inhibitory or activating relationships between transcription factors and their targets center around gene regulatory network (GRN) inference algorithms. Algorithms that utilize high throughput transcriptomic data include information theoretical approaches (e.g. ARACNE, RELNET and CLR), Bayesian networks, and graphical Gaussian models [11]. Capturing dynamic relationships in these networks requires time course transcriptomic data because these data identify intermediate steps leading to final steady state concentrations [11]. Methods that use time course data include dynamic Bayesian networks, linear differential equations, and S-system models [11].

Many GRN inference algorithms depend on extensive prior knowledge to build a TFN network. One of the goals of the CDAA developed by Koryachko et al. was to generate a testable network from microarray data with uneven time points without prior knowledge [24]. This algorithm uses a time lagged approach to compare expression patterns between known iron related targets and putative regulators or influencers of those targets. The CDAA divided the transcriptome time course data (0-72 hrs) into three stages of activity in which each stage was assumed to regulate the next. It was used to predict influential relationships between novel regulators (active in the first stage of expression activity; 0-12 hrs) and known iron response genes (active in the second stage of activity; 12-24 hrs). By aligning changes in expression patterns, network edges between the

putative regulators in the first (initiation) stage and the known iron genes in the second (early response) stage were predicted. The majority of these predicted edges were experimentally validated using qRT-PCR, with one of the predicted edges validated as a direct connection (using Y1H). The regulating TFs identified by this algorithm were previously unknown to be linked to iron homeostasis.

In Lavenus et al., 2012, the Time Delay Correlation (TDCor) algorithm was used to infer the GRN that controls lateral root formation using time-course transcriptomic data of lateral root primordia dissections [38]. This algorithm identifies a set of plausible interactions of a reduced set of genes, culls this set using minimal prior information regarding the role of the gene in the network (regulator or nonregulator, activator or inhibitor), and discerns whether interactions are likely to be direct, indirect or a result of coexpression. They were able to identify 14 out of 23 experimentally validated connections, which was an improvement on the results from two algorithms to which it was compared.

Jaeger et al., identified two network architectures that could explain behavior seen during floral transition [39]. One architecture involved autoregulation of the bZIP transcription factor, FD, while the other involved a feedback loop whereby FD regulates and is regulated by LEAFY (LF). Initially they used a mathematical model to determine which architecture best describes the system behavior. However, the results of the mathematical model suggested that either architecture was plausible. Therefore, promoter reporters and mutant lines were used to determine that the system behavior arose from a feedback loop involving a different gene rather than autoregulation. This study highlights the need to have redundant strategies to explore the properties of the system. In this case, when computational techniques were unable to tease out the nature of the regulation, experimental results were able to determine the type of regulation [39].

3.3. Databases that integrate information

There are a variety of databases that contain known transcription factors and/or *cis*-regulatory elements, which can be used to build TFNs. These databases have been culled from the literature, or derived from global analysis of similar experimental data.

The AthaMap database brings together multidimensional data, [40] and is able to generate a map of predicted binding sites for transcription factors, as well as target sites for small RNAs across the entire Arabidopsis genome. It currently has data on binding sites for 126 transcription factors, which represents 29 different families. AthaMap has been updated to include links to the Arabidopsis Information Portal. It was used to verify downstream targets of the FOREVER

YOUNG FLOWER (FLF) gene involved in floral senescence [41] and to identify novel regulatory sequences by comparing motif families to known cis-regulatory sequences in genes involved in response to pathogen infection [42].

STIFDB2 (an updated version of STIFDB) includes stress response data for three species including Arabidopsis [43]. It also contains information about transcription factors, consensus binding sites, and stress-responsive genes for 15 stress signals and 10 transcription factor families. Binding sites were predicted with the STIF algorithm [44] which searched Hidden Markov Models encoding consensus binding site data. These binding sites are categorized based on gene name, stress, and chromosome. STIFDB2 was used to aid in the identification of G-protein α -subunit (GPA1) and G-protein coupled receptors (GCR1)-responsive genes under abiotic stress [45].

The HRGRN database was recently created as a way to visualize network information across multiple scales, including signal transduction, metabolism and gene regulation [46]. This hormone-specific database brings information together from a diverse set of sources. Node data was derived from TAIR, KEGG, miRBase, and current literature. Edge data was pulled from several databases including TRANSFAC, psRNATarget, Biogrid, AtPINE, KEGG, and PhosPhAt. HRGRN's strength lies in the fact that it uses a scalable graph management database to store biological interactions between genes, proteins and small RNAs, all of which have been confirmed either experimentally or computationally. The utility of this database lies in its categorization of relationships between elements, which facilitates the discovery of previously unknown relationships. Two case studies were completed using HRGRN. One identified ARF12 as a novel link between IAA28 (an auxin-responsive protein) and ATCUL3 (protein binding, ubiquitin-protein) ligase. The other was a subnetwork analysis of WRKY33 (a TF that regulates multiple pathways), AFB1-3 (F-box genes) and TIR1 [46].

Information stored in databases can be used to verify relationships, reduce literature search time, and propose novel relationships between components of interest (e.g, transcription factors and their targets). However, there is no guarantee that the data contained therein will be updated or maintained. Fortunately, researchers have moved toward the development of portals that bring the data together in one location. iPlant (<http://www.iplantcollaborative.org>) is one of these initiatives that focuses on developing a cyber-infrastructure for the plant science community.

3.4. Elucidating emergent properties

Network motifs that lead to emergent properties are often revealed by the elucidation and study of specific transcriptional regulatory networks. Identification and quantification of the components (transcription factors) and interactions (activation or inhibition) of these networks can be used to generate mathematical models that enable the prediction of complex behaviors. Although this approach is challenging in a complex system such as plants, a variety of techniques are being used to generate mathematical models of plant behavior. As described in **Section 4**, systems of ordinary and partial differential equations are most commonly used to mathematically model plant TFN behavior. However, systems of piecewise linear differential equations, Dynamic Boolean networks, and agent-based modeling could also potentially be used [11,24,47-52].

Ordinary and partial differential equations (ODEs and PDEs, respectively) model gene behavior in a continuous fashion. ODEs and PDEs provide detailed quantitative information about the mechanics of a network. However, they require parameter values, the number of which grows exponentially with the number of components being modeled. Such parameters include transcription and translation rates, Hill coefficients, diffusion and degradation rates. Hill equations model enzyme-substrate activity as well as binding affinity/dynamics, while mass action kinetics and diffusion equations are used to model concentration changes in time and space [10]. For well-defined networks, parameters can be derived from existing literature [53]. For small networks, the parameter space may be searched to identify workable ranges for parameters by comparing the output of the model to existing data [39]. Finding parameters for larger networks often requires large amounts of experimental time-course data.

An alternative to general nonlinear ODEs, piecewise linear differential equations (PLDEs), has been used to generate qualitative information from regulatory networks involved in floral morphogenesis [54,55]. Whereas ODEs describe production of components on a continuum, PLDEs describe production in response to an “on” or “off” signal [48]. They are made up of combinations of piecewise linear functions and as such can provide quantitative information about the system in a way that Boolean models cannot. PLDEs are attractive because they provide simplified equations although the number of parameters is comparable to ODEs and the global behavior of these models can be hard to interpret [48]. PLDEs and Boolean dynamic models describe parts of the model in a discrete fashion such that a signal is either on or off, present or not present. This simplification can be problematic in signaling cascades in which activation thresholds play an essential role in the interplay of components [56].

Boolean dynamic systems provide information about final steady states and the sequence of discrete states (trajectories) leading to these final states by treating TFs as inputs that can be turned on or off and TF-target interactions as Boolean

logical operations [11,48]. Emergent properties can be characterized as changes in the number and size of these steady states and the presence of oscillating cycles through a series of discrete states. The advantage of Boolean dynamic systems is that mathematical parameters such as transcription, degradation or diffusion rates, and initial concentrations, are not required to run the system. Additionally, these systems can provide all possible trajectories leading to a set of attractors (steady states) [57]. However, they are limited because they only provide qualitative dynamic behavior. Additionally, the output of the system can be dramatically altered depending on how logical relationships are modified. Despite these limitations, Boolean networks can be a first step to understanding the underlying mechanics of a system. For example, Boolean networks were used to model floral morphogenesis as predicted by the ABC model [58].

Agent based modeling, which has been used to model epidemics and immune function in humans, is based on the principle that complex behavior can arise from simple interactions between groups of individuals [59]. These models incorporate agents as interacting components that operate on a discrete set of basic rules. The difficulty in implementing agent-based models lies in defining the rules by which the agents interact. Agent-based models alleviate the intense parameter requirements of ODEs, but can still provide qualitative information about changes in the system. Although not used extensively, this technique could be a powerful tool for the study of TFNs [60].

ODEs and other types of models provide insight into the mechanisms of network function. However, they must be tested *in silico*, *in vitro* or *in planta* and remodeled in an iterative fashion to generate a final model that predicts plausible outcomes resulting from changes to inputs of the system (typically reflected in changes in transcription factor behavior). Such models can also be examined to identify new relationships and interactions that lead to novel behavior.

4. TRANSCRIPTION FACTOR NETWORKS IN ARABIDOPSIS CONTAINING MOTIFS ASSOCIATED WITH EMERGENT PROPERTIES

The study of emergent properties in complex organisms such as plants is a challenging and developing field because it requires mathematical modeling to elucidate if and how transcriptional regulatory motifs cause these behaviors. However, if a system exhibits a phenotype that is directly tied to certain types of network motifs, it is possible to see evidence of these properties in an experimental setting. An example of this would be the negative feedback loop that controls circadian rhythm in Arabidopsis. In this case, mutation of *TIMING OF CAB EXPRESSION 1* (*TOC1*), a pseudo response regulator involved in the generation of circadian rhythms, exhibits shortened stomatal conductance rhythms compared to wild type plants [61,62].

Additional analysis of motifs is required to associate their behavior with that of an emergent property. Using common biological and computational approaches, researchers have uncovered intricate networks of regulation, in which motifs have been identified. Modeling of the system to determine these properties, however, has been limited. Properly modeling these systems can require several parameters that feed into covariance equations, Boolean Networks analysis, and coupled Ordinary Differential Equations (ODEs) [63]. Here we present advances in our understanding of TFNs, which contain motifs that exhibit emergent properties important for development and stress response.

4.1 Emergent properties in development and cell type maintenance:

Plant root systems are developmentally responsive, altering lateral root and root hair density, and changing or maintaining individual cell types, cell patterning, and other processes in response to their environment [64,65]. The root's ability to readily change and maintain such characteristics is, in part, due to the production of new cells by the stem cell niche. Stem cells produce distinct tissue layers, single-cell files that extend and form the vascular, endodermal, cortex, and epidermal cells [65]. Due to the simplicity of the Arabidopsis root, regulatory processes involved in maintenance of the stem cell niche and root cell patterning are well studied and a number of transcriptional motifs that lead to emergent properties such as bistability, multistability, and dampened reduction, followed by accelerated recovery have been revealed.

4.1.1 Development and Cell Maintenance

Although extensively reviewed, a brief overview of transcriptional regulation involved in maintenance of the stem cell niche and differentiation of stem cell initials and their respective cell layers is warranted [47,65,66]. Two nested feedback circuits that include SHORT ROOT (SHR), SCARECROW (SCR), RETINOBLASTOMA-RELATED (RBR), and CYCLIN D6 (CYCD6;1), along with an auxin gradient results in precise spatial confinement of asymmetric cell division (ACDs), switching the system "on". The bistable switch is turned off when increased protein degradation occurs [47,66]. These studies are excellent examples of how one could identify motifs, mathematically model systems, and biologically test the model. It should be noted that these studies involved an iterative approach to modeling where ultimately critical parameters were identified that could explain plant behavior at the molecular and developmental level.

Beyond the root stem cell niche, cellular lineage is highly regulated and dependent on proper ACDs. ACD at the cortex/endodermal initial (CEI) in the

stem cell niche gives rise to ground tissue (cortex and endodermal cell layers). Anticlinal division of CEI cells produce the CEI daughter (CEID) cells and periclinal division generates endodermal and cortex precursors [67,68]. There are two identified networks controlling ground tissue formation. One network regulates cell lineage and a second regulates cellular division and patterning [69,70]. Although little is known about motifs associated with cortex formation, we can apply logical gate theory to transcriptional control for anticlinal and periclinal cell divisions that produce the endodermis lineage [69]. There exists a logical AND relationship between JACKDAW (JKD) and its close homolog BALDIBIS (BIB), which activates SCR. SHR is prevented from exiting the nucleus and promotes an additional logical AND relationship between SHR-SCR. The promotion of the second logical AND relationship of SHR-SCR leads to a positive feedback loop, where several downstream genes are activated, one of which is JKD, a regulator of SCR. In addition to JKD, MAGPIE (MGP) and NUTCRACKER (NUC) are downstream targets of SHR-SCR, where their activation leads to periclinal cell division that generates the cortex and endodermis, as well as specification for endodermal cell fate [69]. Thus, the activation of SHR-SCR requires the JKD-BIB complex, and the JKD activity is modulated by SHR-SCR (Figure 4a). In a dynamic nonlinear system, this kind of positive feedback loop where the SHR-SCR complex activates the regulator of SCR possibly leads to multistability and switch-like behavior of transcript [14,71]. Such behavior may be a common feature of cell division [16,17,47,66].

Transcriptional motifs described in **Section 2** have also been identified and modeled in longitudinal root development. There are three distinct developmental phases of root cells: the meristematic zone, where cells actively divide to generate a pool of cells that will elongate and differentiate; the elongation zone, where cells lose their ability to divide and increase in length; and, the differentiation zone, where cells differentiate and acquire their specialized characteristics and functions. There are two sub-stages: the transition zone, which falls between the meristematic zone and the elongation zone; and the early differential zone, which falls between the elongation and differential zone [65,72]. Cellular differentiation in longitudinal root development involves an incoherent feed-forward type II motif, where PHABULOSA (PHB), a HD-ZIPIII transcription factor, directly activates the cytokinin (CK) biosynthesis gene ISOPENTENYL TRANSFERASE 7 (IPT7) in the meristematic zone of the root. This activation leads to RESPONSE REGULATOR (ARR1) activation in the early differentiation zone. ARR1 represses vascular expression of PHB in the transition zone and represses microRNA165. An additional feedback loop also occurs, in which CK production represses PHB expression to sustain the balance of cell division and differentiation during root development. Modeling this system using the Runge-Kutta method of ODEs found that CK inhibition of mir165 transcription

results in a dampened reduction, and accelerated recovery of PHB in response to a temporary CK stimulus [50].

Key networks have also been identified in lateral root formation. Lateral roots arise from asymmetric anticlinal divisions of xylem pole pericycle cells. These divisions lead to anticlinal and periclinal divisions of lateral root primordia cells [73,74]. Using the Time Delay Correlation (TDCor) algorithm to infer the GRN that controls lateral root formation, Lavenus et al., were able to create a coarse network topology linking early activated genes involved in lateral root initiation, with initiation of a regulatory cascade. They also found that late activated genes involved in lateral root cell identity exhibited reduced organization of network topology and had a high level of positive and negative feedback loops [38]. Interestingly, they found several meristematic genes in the late activated gene network. It is likely that the abundance of meristematic genes causes the high density of positive feedback loops [38], which can cause gene oscillation when they are interconnected with negative feedback loops [5].

Regulatory motifs associated with emergent behaviors have also been shown to play essential roles in floral development. A recent study examining floral transition investigated the dynamics of several interlocking activity hubs within a GRN. Through the interlocking connections, feedback and feed-forward loops connected key genes associated with floral transition [39]. Although individual hubs involved in this process were well known, the study sought to uncover a more complete set of interactions and explore the dynamics of the GRN network. Individual activity hubs, grouped by their controlling transcription factor (i.e. FLOWERING LOCUS T (FT), LEAFY (LFY), TERMINAL FLOWER1 (TFL1), APETALA1 (AP1), and bZIP transcription factor (FD)), were modeled with ODEs using Hill-type gene activation and repression. At the center of this interconnected GRN is a core feed-forward loop that exhibits irreversible behavior and has the ability to filter noise that regulates floral transition [39,75]. Although the mathematical model incorporated network simplification, this approach was unique and offers insight into alternative ways of modeling systems that contain various gene activity hubs. Simplifying the model by grouping interacting genes into activity hubs imparts the big picture and phenotype associated with the system being studied. However, oversimplification can also lead to various vague conclusions, as mentioned in **Section 3.2**. Considering GRNs as interlocking activity hubs rather than singular hubs containing regulation and regulatory check (feedbacks) may expand our understanding of dynamic processes with antagonistic, competing or crosstalk characteristics, such as hormone signaling, the interplay between iron and phosphate, or the crosstalk between zinc and auxin [76-78].

4.2 Emergent properties in abiotic stress response

Plants rapidly change transcript activity to adjust to their environment. These transcriptional changes bring about molecular and biochemical adjustments, and phenotypic adaptations. Below we provide examples of TFNs involved in abiotic stress response. Essential to these TFNs are motifs that are associated with emergent properties, leading to distinct transcript activities.

4.2.1 Drought

Plants initialize many responses to drought stress. One strategy includes escaping drought through early flowering and reducing the size of the plant to increase water use efficiency. Bechtold et al., identified AGAMOUS-LIKE22 (ATAGL22), a transcription factor that represses flower development and affects primary metabolism and developmental processes, as a central hub gene involved in drought stress GRN [52]. Drought response was linked to development through Metropolis Variational Bayesian State Space Modeling (M-VBSSM). M-VBSSM coupled with mutant transcript analysis, identified an intricate network of transcription factors involved in modulating plant response to drought. Based on the network developed, several motifs could potentially be mathematically modeled to identify complex behavior. Mutant analysis indicated that the bHLH TF, bHLH038 auto activates itself presumably in response to drought and represses AT1G55760, a BTB/POZ domain-containing protein, in addition to AGL22, constituting a coherent type II feed-forward loop. Additionally, AT1G55760 activates itself to activate AGL22, which repress bHLH038. bHLH038 represses AT1G55760, thus, these double inhibitory relationships constitute a positive feedback motif [52] (Figure 4b). These motifs can be embedded within each other and function simultaneously or separately. Presumably, the positive auto-regulations, and the positive feedback loops lead to bistability [5]. The next step is to use a more advanced modeling approach, such as ODEs, to describe the dynamics of genes involved in this process, and to determine the conditions for each motif. Such an effort could provide insight into ways to mitigate drought stress at the transcriptional level.

4.2.2 Cold

Regulatory motifs associated with cold acclimation have also been identified. Low temperature-induced alterations in gene expression provoke a wide array of physiological and biochemical adjustments that enable plants to acclimate to lower temperatures. We can apply Boolean logic to the network controlling cold signaling, whereby ICE1 or ICE2 activates ICE-C-REPEAT BINDING FACTORS 1,2,3 (CBFs). The activation of CBF2 indirectly triggers the repression of CBF3, constituting a transcriptional feedback loop that attenuates cold signaling and stomatal development (Figure 4c) [79]. During acclimation such a feedback motif

could bring the system back to a steady state transcript concentration in response to chaotic behavior triggered by cold stress.

4.3 Emergent properties involved in nutritional stress

Transcription factor networks involved in nutritional stress response contain motifs that facilitate emergent properties important for maintaining healthy levels of nutrients in the face of changing nutrient availability. Networks that control homeostasis of nutrients such as copper, iron and zinc are tightly regulated because, while they are essential for cellular processes such as photosynthesis and respiration, in excess they cause metal toxicity [80,81]. In addition, copper and iron deficiency alters the amplitude of oscillation of key circadian clock genes [82,83].

4.3.1 Copper:

In response to copper starvation ELONGATED HYPOCOTYL5 (HY5), a circadian clock gene, and SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 7 (SPL7), a constitutively expressed TF that up-regulates COPPER TRANSPORTER (COPT2 and COPT6) co-regulate miR408 to maintain copper homeostasis [84,85]. Thus, cross talk between light and copper leads to an incoherent type I feed-forward loop in which the HY5 and SPL7 logical AND relationship activates mir408 to repress LAC12 while HY5 directly activates LAC12. In addition, HY5 directly represses SPL7, which, in turn, represses HY5 expression, constituting a positive feedback motif. Notably, HY5, SPL7, and mir408 exhibit pulse-like gene expression with peaks dependent on time [85]. Thus, the incoherent type I feed-forward loop associated with this network may be responsible for weak pulse-like and accelerated expression of HY5 whereas the positive feedback motif of HY5 and SPL7 could lead to multistable states [5]. Modeling this gene regulatory network using ODEs will provide insight into whether the pulse-like expression is solely due to circadian effects, the incoherent type I feedback loop, or a combination of both.

4.3.2 Iron:

Under iron deficient conditions, several early active transcription factors (MYB DOMAIN PROTEIN55 (MYB55), ARABIDOPSIS 6B-INTERACTING PROTEIN 1-LIKE 2 (ASIL2), EARLY TRANSCRIPTION FACTOR (ETF9), and CONSTANS-like (COL4)) have been shown to modulate known members of the iron deficiency transcriptional cascade [24]. These early responsive genes appear to be involved in several motifs of interest (Figure 4c-d). ASIL2 indirectly targets and modulates the expression of BRUTUS (BTS), POPEYE (PYE), group IVc bHLH BASIC HELIX-LOOP-HELIXs (bHLHs), and MYB DOMAIN PROTEINs

(MYBs). Based upon these findings ASIL2 may initiate a coherent type II feed-forward loop that includes a logical AND relationship between group IVc bHLH family members to activate MYB family members and other downstream targets that positively regulate iron acquisition. Separately, ASIL2 may initiate another coherent type II feed-forward loop that includes the repression of BTS, an E3 ligase that degrades group IVc bHLH family members. In literature, these same bHLH family members have been shown to have either logical AND or OR relationships that activate PYE, which directly represses several targets to negatively regulate iron acquisition [37,63,85-88]. Since BTS expression is increased in a *pye* mutant, PYE is also involved in a feedback loop indirectly repressing BTS. In addition, a bHLH target degraded by BTS, bHLH105 (ILR3) autoregulates itself. Perhaps ILR3 autoregulation acts to offset protein degradation by BTS, leading to the activation of PYE (Figure 4d-e) [37,63,85-88].

Although not highlighted here, the early-activated TFs identified in Koryachko et al confer both positive or negative regulation on shared targets [24]. For this reason, studying combinatorial effects of multiple TFs on specific targets is essential for modeling iron homeostasis. Also, the iron-responsive E3 ligases, BTS and RGLG1/2 (not discussed in this network) function to degrade their protein targets, as such, parameters for modeling iron homeostasis should account for protein decay [89].

4.3.3 Zinc:

Comparative microarray analysis of the root transcriptome in Arabidopsis wild type and *bZIP19* and *bZIP23* (*m19m23*) double mutant plants, identified several differentially expressed genes, many of which contained the zinc deficiency response element (ZDRE) within their promoters, including ZINC TRANSPORTER PRECURSORS (ZIPs). Direct binding of these bZIP TFs to the promoters of ZIP transporters was confirmed using Y1H analysis. Thus, zinc homeostasis is thought to be transcriptionally regulated by two BASIC-REGION LEUCINE ZIPPER family members, bZIP19 and bZIP23, which regulates zinc transporters, ZIP1/3/4/5/9/10/12 [90].

Using a system of four ODEs, zinc homeostasis was modeled taking into account zinc internal and external concentration, expression of genes that encode zinc transporters, transcription factors that dimerize to activate the zinc transporters, and transcription factor that inhibits the activators. Whereas component one of the system of ODEs modeled changes in transporter gene activity, component two modeled changes in internal zinc concentration, component three modeled changes in dimerization activity, and component four modeled changes in putative transcriptional inhibitors. This approach revealed that internal zinc is a function of transporter activity and external zinc concentration, while transporter

activity is controlled by the dimerization of transcriptional activators, bZIP19 and bZIP23. Mathematically, this activation is a function of activator/inhibitor binding and inhibition is a function of internal zinc concentration and activator/inhibitor binding. Thus, this comprehensive ODE model helped elucidate a zinc regulatory system that contained a feedback loop in which internal zinc concentration initiated the repression of the activators. Additionally, in response to external zinc concentration, two bifurcations events were revealed that lead to oscillation of internal zinc concentration and ZIP transporters [49,51].

It should be noted that rather than taking into account the effects of each of the seven ZIP transporters targeted by bZIP19 and bZIP23 this modeling approach used a summation of ZIP transporter activities. In addition, this approach does not take into account additional pathways involved in zinc homeostasis. Despite these simplifications, this study successfully revealed the relationship between internal and external zinc supply and zinc responsive TFs and their targets. Thus, this study sets precedence in the area of transcriptional control of metal stress whereby a plausible TFN identified emergent behavior using a computational approach.

The emergence of feedback loops and bifurcations in metal homeostasis is not surprising. In electrical systems, bifurcations are useful when a system needs a low and high state. In the case of deficiencies in zinc and in many other metals, plants have to quickly acquire metal when they sense low internal concentration. However this process can lead to excess metal if not properly regulated, causing the system to over-accumulate zinc. Thus metal deficiency responses may often correspond with feedback loops, which are associated with bringing a system back to steady state. In a simple scenario, if plants experience a situation where they are zinc deficient (low state), transcriptional control kicks into overdrive, thereby taking in excess zinc (high state). This transition from low to high state would cause the system to eventually accumulate too much zinc, requiring an inhibitory signal to stop the transport of zinc (hence, the feedback loop, which brings the system back to steady state).

Although not discussed at length in this article, a recent review modeled the behavior of copper transporters in response to cytosolic copper ions concentration. Using Michaelis-Menten kinetics and coupled ODEs, Penarrubia et al., concluded that a copper autoregulatory negative feedback loop leads to oscillation of copper concentrations, COPPER TRANSPORTATION (COPT) mRNA and COPT protein over time [91]. In the future, advances in generating predictive TFNs models involved in nutrient stress response will require analysis of the effects of other key parameters, such as internal and external metal ion concentration, on the system.

5. CONCLUSIONS

The responses of plants to abiotic and biotic stresses are driven by complex biochemical cascades, which ultimately allow plants to effectively adapt to changing environments. Transcription factors and the regulatory networks they control serve as key components in these responses. Understanding the emergent properties seen in TFNs is a critical step towards the engineering of novel plants that can withstand abiotic stress and sustain robust growth. Characterizing the network motifs that give rise to these properties seen in TFNs will require the formulation of mathematical models that capture their dynamic behaviors. In this review, we outlined the experimental and computational techniques that have been used in the literature to move toward the ultimate goal of quantifying emergent properties of TFNs in Arabidopsis.

Decades of research have led to the static description of transcriptional motifs, outlining the activity and impact of individual TFN components. Characterizing bistability, sustained oscillations, hysteresis, and homeostasis require a framework for quantitatively analyzing the dynamic behavior of transcriptional motifs individually or in combination. High resolution dynamic transcriptional, posttranscriptional, and posttranslational data across time and space provides the information that allows for the effective screening of putative parameters involved in specific TFN responses. Y1H and ChIP-SEQ are effective experimental approaches for identifying specific transcription factors and their targets. GRN inference algorithms provide computational tools that elucidate the direction of influence between components in the network. The integration of curated databases yields useful information for refining and validating new networks. Model structures such as ordinary differential equations, piecewise linear differential equations, dynamic Boolean networks, and agent-based models provide the mathematical framework needed to quantify the properties associated with specific networks. These approaches, individually or combined, have been shown to be successful in a few targeted examples. The broader application of these techniques, however, will require further development of novel experimental and modeling approaches. Several of the systems reviewed here utilize model simplification, by grouping components or using simplifying assumptions and/or models. Examples include studies of floral development and zinc transport, which, rather than focusing on individual genes, grouped them into hubs [39, 49, 51]. These simplifications are necessary to reduce model complexity or deal with unknown aspects of the system such as model parameters. However, simplifications can make validation more difficult and keep salient details of regulation from discovery. As additional information is discovered about the biological system, increasing levels of complexity can be introduced into the model. In this way, we can learn about the behavior of increasingly complex interactions.

As we move towards generating more advanced algorithms that accurately predict plant behavior new tools are being developed to extract more information from increasingly large datasets, enabling better system estimation. New bioinformatics tools are needed to capture relevant patterns from high-resolution dynamic transcriptional data. To attain more accurate information about how TFNs are regulated, post-transcriptional events such as microRNAs processing and ribonuclease activity, and post-translational modifications such as phosphorylation and protein degradation rates across time and space will need to be incorporated into models. A deeper investigation into the nature of complex motif interactions is also required. In the way that we have learned about emergent properties associated with three-component motifs [5], we will also need to identify trends associated with nested or linked motifs. Finally, many other salient features, including protein movement, metabolic flux, and nutrient content are critical parameters that must be considered while elucidating TFNs. Fortunately, researchers are increasingly merging mathematical tools from diverse fields such as systems biology, engineering, and bioinformatics with those from molecular biology, physiology and biochemistry. These efforts will enable the generation of models with increasing resolution and predictive power that generate *in silico* data upon which biologists can more comfortably rely, reducing the need for intensive validation

ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. MCB 1247427. We thank Anna Matthiadis for critical reading of this manuscript.

REFERENCES

- [1] A. Aderem, Systems Biology: Its Practice and Challenges, *Cell*. 121 (2005) 511–513. doi:10.1016/j.cell.2005.04.020.
- [2] V. Reinke, Transcriptional regulation of gene expression in *C. elegans*, *WormBook*. (2013) 1–31. doi:10.1895/wormbook.1.45.2.
- [3] R. Pinho, V. Garcia, M. Irimia, M.W. Feldman, Stability Depends on Positive Autoregulation in Boolean Gene Regulatory Networks, *PLoS Comput. Biol.* 10 (2014) e1003916. doi:10.1371/journal.pcbi.1003916.
- [4] E. Wong, B. Baur, S. Quader, C.H. Huang, Biological network motif detection: principles and practice, *Briefings in Bioinformatics*. 13 (2012) 202–215. doi:10.1093/bib/bbr033.
- [5] U. Alon, Network motifs: theory and experimental approaches, *Nat Rev Genet.* 8 (2007) 450–461. doi:10.1038/nrg2102.
- [6] J.J. Tyson, K.C. Chen, B. Novak, Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell, *Current Opinion in Cell Biology*. 15 (2003) 221–231. doi:10.1016/S0955-0674(03)00017-6.
- [7] J.K. Parrish, L. Edelstein-Keshet, Complexity, Pattern, and Evolutionary Trade-Offs in Animal Aggregation, *Science*. 284 (1999) 99–101. doi:10.1126/science.284.5411.99.
- [8] A. Cruz-Ramírez, S. Díaz-Triviño, I. Blilou, V.A. Grieneisen, R. Sozzani, C. Zamioudis, et al., A Bistable Circuit Involving SCARECROW-RETINOBLASTOMA Integrates Cues to Inform Asymmetric Stem Cell Division, *Cell*. 150 (2012) 1002–1015. doi:10.1016/j.cell.2012.07.017.
- [9] Y. Li, K. Varala, G.M. Coruzzi, From milliseconds to lifetimes: tracking the dynamic behavior of transcription factors in gene networks, *Trends in Genetics*. 31 (2015) 509–515. doi:10.1016/j.tig.2015.05.005.
- [10] G. Karlebach, R. Shamir, Modelling and analysis of gene regulatory networks, *Nat Rev Mol Cell Biol.* 9 (2008) 770–780. doi:10.1038/nrm2503.
- [11] M. Hecker, S. Lambeck, S. Toepfer, E. van Someren, R. Guthke, Gene regulatory network inference: Data integration in dynamic models—A review, *Biosystems*. 96 (2009) 86–103. doi:10.1016/j.biosystems.2008.12.004.
- [12] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Network motifs: simple building blocks of complex networks, *Science*. 298 (2002) 824–827.
- [13] A.M. Sullivan, A.A. Arsovski, J. Lempe, K.L. Bubb, M.T. Weirauch, P.J. Sabo, et al., Mapping and Dynamics of Regulatory DNA and Transcription Factor Networks in *A. thaliana*, *Cell Reports*. 8 (2014)

2015–2030. doi:10.1016/j.celrep.2014.08.019.

[14] D. Angeli, J.E. Ferrell, E.D. Sontag, Detection of multistability, bifurcations, and hysteresis in a large class of biological positive-feedback systems, *Proceedings of the National Academy of Sciences*. 101 (2004) 1822–1827. doi:10.1088/1478-3975/12/4/045007/meta;jsessionid=CA00F88AF8EA420A9738D9BAA092178B.c2'].
 [15] E. Hopf, Bifurcation of a periodic solution from a stationary solution of a system of differential equations, *Berlin Mathematische Physics Klasse, Sachsischen Akademik Der Wissenschaften Leipzig*. 94 (1942) 3–32.
 [16] M. Laurent, N. Kellershohn, Multistability: a major means of differentiation and evolution in biological systems, *Trends in Biochemical Sciences*. 24 (1999) 418–422. doi:10.1016/S0968-0004(99)01473-5.
 [17] E. Azpeitia, M. Benítez, I. Vega, C. Villarreal, E.R. Alvarez-Buylla, Single-cell and coupled GRN models of cell patterning in the *Arabidopsis thaliana* root stem cell niche, *BMC Systems Biology* 2010 4:1. 4 (2010) 1. doi:10.1186/1752-0509-4-134.
 [18] J.S. Hallinan, P.T. Jackway, *Network Motifs, Feedback Loops and the Dynamics of Genetic Regulatory Networks*, IEEE, 2005. doi:10.1109/CIBCB.2005.1594903.
 [19] L. Yan, Z.-Q. Liu, Y.-H. Xu, K. Lu, X.-F. Wang, D.-P. Zhang, Auto- and Cross-repression of Three *Arabidopsis* WRKY Transcription Factors WRKY18, WRKY40, and WRKY60 Negatively Involved in ABA Signaling, *J Plant Growth Regul.* 32 (2012) 399–416. doi:10.1007/s00344-012-9310-8.
 [20] M. Nuruzzamen, A.M. Sharoni, S. Kikuchi, Roles of NAC transcription factors in the regulation of biotic and abiotic stress responses in plants, (2013) 1–16. doi:10.3389/fmicb.2013.00248/abstract.
 [21] H. Meng, Y. Wang, Cis-acting regulatory elements: from random screening to quantitative design, *Quant Biol.* 3 (2015) 107–114. doi:10.1007/s40484-015-0050-1.
 [22] L. Taher, L. Narlikar, I. Ovcharenko, Identification and Computational Analysis of Gene Regulatory Elements, *Cold Spring Harb Protoc.* 2015 (2015) pdb.top083642. doi:10.1101/pdb.top083642.
 [23] J. William A Schmitt, R.M. Raab, G. Stephanopoulos, Elucidation of Gene Interaction Networks Through Time-Lagged Correlation Analysis of Transcriptional Data, *Genome Research*. 14 (2004) 1654–1663. doi:10.1101/gr.2439804.
 [24] A. Koryachko, A. Matthiadis, D. Muhammad, J. Foret, S.M. Brady, J.J. Ducoste, et al., Clustering and Differential Alignment Algorithm: Identification of Early Stage Regulators in the *Arabidopsis thaliana* Iron Deficiency Response, *PLoS ONE*. 10 (2015) e0136591–21. doi:10.1371/journal.pone.0136591.

- 972 [25] P. Ouwerkerk, A.H. Meijer, Yeast One-Hybrid Screening for DNA-Protein
973 Interactions - Current Protocols in Molecular Biology - Ouwerkerk - Wiley
974 Online Library, Current Protocols in Molecular (2001).
975 doi:10.1002/0471142727.mb1212s55/full"].
- 976 [26] N. Mitsuda, M. Ikeda, S. Takada, Y. Takiguchi, Y. Kondou, T. Yoshizumi,
977 et al., Efficient Yeast One-/Two-Hybrid Screening Using a Library
978 Composed Only of Transcription Factors in *Arabidopsis thaliana*, *Plant*
979 *Cell Physiol.* 51 (2010) 2145–2151. doi:10.1093/pcp/pcq161.
- 980 [27] A. Gaudinier, L. Zhang, J.S. Reece-Hoyes, M. Taylor-Teeple, L. Pu, Z.
981 Liu, et al., Enhanced Y1H assays for *Arabidopsis*, *Nature Methods.* 8
982 (2011) 1053–1055. doi:10.1038/nmeth.1750.
- 983 [28] J.L. Pruneda-Paz, G. Breton, D.H. Nagel, S.E. Kang, K. Bonaldi, C.J.
984 Doherty, et al., A Genome-Scale Resource for the Functional
985 Characterization of *Arabidopsis* Transcription Factors, *CellReports.* 8
986 (2014) 622–632. doi:10.1016/j.celrep.2014.06.033.
- 987 [29] M.L. Bulyk, Protein Binding Microarrays for the Characterization of DNA–
988 Protein Interactions, in: *Analytics of Protein–DNA Interactions*, Springer
989 Berlin Heidelberg, 2006: pp. 65–85. doi:10.1007/10_025.
- 990 [30] R.E. Thurman, E. Rynes, R. Humbert, J. Vierstra, M.T. Maurano, E.
991 Haugen, et al., The accessible chromatin landscape of the human
992 genome, *Nature.* 489 (2012) 75–82. doi:10.1038/nature11232.
- 993 [31] S. Neph, J. Vierstra, A.B. Stergachis, A.P. Reynolds, E. Haugen, B.
994 Vernot, et al., An expansive human regulatory lexicon encoded in
995 transcription factor footprints, *Nature.* 489 (2012) 83–90.
996 doi:10.1038/nature11212.
- 997 [32] K. Kaufmann, J.M. Muiño, M. Østerås, L. Farinelli, P. Krajewski, G.C.
998 Angenent, Chromatin immunoprecipitation (ChIP) of plant transcription
999 factors followed by sequencing (ChIP-SEQ) or hybridization to whole
1000 genome arrays (ChIP-CHIP), *Nature Protocols.* 5 (2010) 457–472.
1001 doi:10.1038/nprot.2009.244.
- 1002 [33] M. Tompa, N. Li, T.L. Bailey, G.M. Church, B. De Moor, E. Eskin, et al.,
1003 Assessing computational tools for the discovery of transcription factor
1004 binding sites, *Nat Biotechnol.* 23 (2005) 137–144. doi:10.1038/nbt1053.
- 1005 [34] T.R. O'Connor, C. Dyreson, J.J. Wyrick, Athena: a resource for rapid
1006 visualization and systematic analysis of *Arabidopsis* promoter sequences,
1007 *Bioinformatics.* 21 (2005) 4411–4413. doi:10.1093/bioinformatics/bti714.
- 1008 [35] K. Toufighi, S.M. Brady, R. Austin, E. Ly, N.J. Provart, The Botany Array
1009 Resource: e-Northern, Expression Angling, and promoter analyses, *The*
1010 *Plant Journal.* 43 (2005) 153–163. doi:10.1111/j.1365-
1011 313X.2005.02437.x.
- 1012 [36] T.L. Bailey, M. Boden, F.A. Buske, M. Frith, C.E. Grant, L. Clementi, et
1013 al., MEME Suite: tools for motif discovery and searching, *Nucleic Acids*
1014 *Research.* (2009) gkp335. doi:10.1093/nar/gkp335.

- 1015 [37] B. Li, A. Gaudinier, M. Tang, M. Taylor-Teeples, N.T. Nham, C. Ghaffari,
1016 et al., Promoter-Based Integration in Plant Defense Regulation, *Plant*
1017 *Physiology*. 166 (2014) 1803–1820. doi:10.1104/pp.114.248716.
- 1018 [38] J. Lavenus, T. Goh, S. Guyomarc'h, K. Hill, M. Lucas, U. Voß, et al.,
1019 Inference of the Arabidopsis Lateral Root Gene Regulatory Network
1020 Suggests a Bifurcation Mechanism That Defines Primordia Flanking and
1021 Central Zones, *The Plant Cell*. 27 (2015) 1368–1388.
1022 doi:10.1105/tpc.114.132993.
- 1023 [39] K.E. Jaeger, N. Pullen, S. Lamzin, R.J. Morris, P.A. Wigge, Interlocking
1024 Feedback Loops Govern the Dynamic Behavior of the Floral Transition in
1025 Arabidopsis, *The Plant Cell*. 25 (2013) 820–833.
1026 doi:10.1105/tpc.113.109355.
- 1027 [40] R. Hehl, L. Norval, A. Romanov, L. Bülow, Boosting AthaMap Database
1028 Content with Data from Protein Binding Microarrays, *Plant Cell Physiol*.
1029 57 (2016) e4–e4. doi:10.1093/pcp/pcv156.
- 1030 [41] B.-L. Chen, L.-Z. Liu, F.-X. Wu, Inferring gene regulatory networks from
1031 multiple time course gene expression datasets, in: 2011: pp. 12–17.
1032 doi:10.1109/ISB.2011.6033114.
- 1033 [42] F. Machens, M. Becker, F. Umrath, R. Hehl, Identification of a novel type
1034 of WRKY transcription factor binding site in elicitor-responsive cis-
1035 sequences from Arabidopsis thaliana, *Plant Molecular Biology*. 84 (2013)
1036 371–385. doi:10.1007/s11103-013-0136-y.
- 1037 [43] M. Naika, K. Shameer, O.K. Mathew, R. Gowda, R. Sowdhamini,
1038 STIFDB2: An Updated Version of Plant Stress-Responsive Transcription
1039 Factor DataBase with Additional Stress Signals, Stress-Responsive
1040 Transcription Factor Binding Sites and Stress-Responsive Genes in
1041 Arabidopsis and Rice, *Plant Cell Physiol*. 54 (2013) e8–e8.
1042 doi:10.1093/pcp/pcs185.
- 1043 [44] A.S. Sundar, S.M. Varghese, K. Shameer, N. Karaba, M. Udayakumar, R.
1044 Sowdhamini, STIF: identification of stress-upregulated transcription factor
1045 binding sites in Arabidopsis thaliana, *Bioinformation*. 2 (2008) 431–437.
- 1046 [45] N. Chakraborty, N. Singh, K. Kaur, N. Raghuram, G-protein Signaling
1047 Components GCR1 and GPA1 Mediate Responses to Multiple Abiotic
1048 Stresses in Arabidopsis, *Front. Plant Sci*. 6 (2015) 1000.
- 1049 [46] X. Dai, J. Li, T. Liu, P.X. Zhao, HRGRN: A Graph Search-Empowered
1050 Integrative Database of Arabidopsis Signaling Transduction, Metabolism
1051 and Gene Regulation Networks, *Plant Cell Physiol*. 57 (2016) e12–e12.
1052 doi:10.1093/pcp/pcv200.
- 1053 [47] A. Cruz-Ramírez, S. Díaz-Triviño, G. Wachsman, Y. Du, M. Arteaga-
1054 Vázquez, H. Zhang, et al., A SCARECROW-RETINOBLASTOMA Protein
1055 Network Controls Protective Quiescence in the Arabidopsis Root Stem
1056 Cell Organizer, *PLOS Biol*. 11 (2013) e1001724.
1057 doi:10.1371/journal.pbio.1001724.

- 1058 [48] H. de Jong, Modeling and Simulation of Genetic Regulatory Systems: A
 1059 Literature Review, *Journal of Computational Biology*. 9 (2002) 67–103.
- 1060 [49] J. Claus, A. Chavarría-Krauser, Modeling Regulation of Zinc Uptake via
 1061 ZIP Transporters in Yeast and Plant Roots, *PLoS ONE*. 7 (2012) e37193–
 1062 11. doi:10.1371/journal.pone.0037193.
- 1063 [50] R. Dello Ioio, C. Galinha, A.G. Fletcher, S.P. Grigg, A. Molnar, V.
 1064 Willemssen, et al., A PHABULOSA/Cytokinin Feedback Loop Controls
 1065 Root Growth in Arabidopsis, *Current Biology*. 22 (2012) 1699–1704.
 1066 doi:10.1016/j.cub.2012.07.005.
- 1067 [51] J. Claus, M. Ptashnyk, A. Bohmann, A. Chavarría-Krauser, Global Hopf
 1068 bifurcation in the ZIP regulatory system, *Journal of Mathematical Biology*.
 1069 (2016) 1–22. doi:10.1007/s00285-014-0836-1.
- 1070 [52] U. Bechtold, C.A. Penfold, D.J. Jenkins, R. Legaie, J.D. Moore, T.
 1071 Lawson, et al., Time-Series Transcriptomics Reveals That AGAMOUS-
 1072 LIKE2 Affects Primary Metabolism and Developmental Processes in
 1073 Drought-Stressed Arabidopsis, *The Plant Cell*. 28 (2016) 345–366.
 1074 doi:10.1105/tpc.15.00910.
- 1075 [53] L.R. Band, J.A. Fozard, C. Godin, O.E. JENSEN, T. Pridmore, M.J.
 1076 Bennett, et al., Multiscale Systems Analysis of Root Growth and
 1077 Development: Modeling Beyond the Network and Cellular Scales, *The*
 1078 *Plant Cell*. 24 (2012) 3892–3906. doi:10.1105/tpc.112.101550.
- 1079 [54] J. Gebert, N. Radde, G.-W. Weber, Modeling gene regulatory networks
 1080 with piecewise linear differential equations, *European Journal of*
 1081 *Operational Research*. 181 (2007) 1148–1165.
- 1082 [55] T.J. Perkins, M. Hallett, L. Glass, Inferring models of gene expression
 1083 dynamics, *Journal of Theoretical Biology*. 230 (2004) 289–299.
 1084 doi:10.1016/j.jtbi.2004.05.022.
- 1085 [56] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts, et al.,
 1086 *Molecular Biology of the Cell*, Sixth Edition:, Taylor & Francis Group,
 1087 2014.
- 1088 [57] R. Albert, J. Thakar, Boolean modeling: a logic-based dynamic approach
 1089 for understanding signaling and regulatory networks and for making
 1090 useful predictions - Albert - 2014 - *Wiley Interdisciplinary Reviews:*
 1091 *Systems Biology and Medicine* - Wiley Online Library, Wiley
 1092 *Interdisciplinary Reviews: Systems* (2014).
 1093 doi:10.1002/wsbm.1273/full].
- 1094 [58] L. Mendoza, E.R. Alvarez-Buylla, Dynamics of the Genetic Regulatory
 1095 Network for Arabidopsis thaliana Flower Morphogenesis, *Journal of*
 1096 *Theoretical Biology*. 193 (1998) 307–319. doi:10.1006/jtbi.1998.0701.
- 1097 [59] C.M. Macal, M.J. North, Tutorial on agent-based modelling and
 1098 simulation, *Journal of Simulation*. 4 (2010) 151–162.
 1099 doi:10.1057/jos.2010.3.
- 1100 [60] V. Vallurupalli, C. Purdy, Agent-based modeling and simulation of

1101 biomolecular reactions, Scalable Computing: Practice and Experience. 8
1102 (2001).

1103 [61] D.E. Somers, A.A. Webb, M. Pearson, S.A. Kay, The short-period mutant,
1104 *toc1-1*, alters circadian clock regulation of multiple outputs throughout
1105 development in *Arabidopsis thaliana*, *Development*. 125 (1998) 485–494.
1106 doi:10.1139/b69-292.

1107 [62] P. Más, D. Alabadí, M.J. Yanovsky, T. Oyama, S.A. Kay, Dual Role of
1108 TOC1 in the Control of Circadian and Photomorphogenic Responses in
1109 *Arabidopsis*, *The Plant Cell*. 15 (2003) 223–236. doi:10.1105/tpc.006734.

1110 [63] A. Koryachko, A. Matthiadis, J.J. Ducoste, J. Tuck, T.A. Long, C.
1111 Williams, Computational approaches to identify regulators of plant stress
1112 response using high-throughput gene expression data, *Current Plant*
1113 *Biology*. (2015) 1–10. doi:10.1016/j.cpb.2015.04.001.

1114 [64] S. Smith, I. De Smet, Root system architecture: insights from *Arabidopsis*
1115 and cereal crops, *Phil. Trans. R. Soc. B*. 367 (2012) 1441–1452.
1116 doi:10.1098/rstb.2011.0234.

1117 [65] J.J. Petricka, C.M. Winter, P.N. Benfey, Control of *Arabidopsis* Root
1118 Development, [Http://Dx.Doi.org/10.1146/Annurev-Arplant-042811-](http://Dx.Doi.org/10.1146/Annurev-Arplant-042811-105501)
1119 [105501](http://Dx.Doi.org/10.1146/Annurev-Arplant-042811-105501). 63 (2012) 563–590. doi:10.1146/annurev-arplant-042811-
1120 [105501](http://Dx.Doi.org/10.1146/Annurev-Arplant-042811-105501).

1121 [66] A. Cruz-Ramírez, S. Díaz-Triviño, I. Blilou, V.A. Grieneisen, R. Sozzani,
1122 C. Zamioudis, et al., A Bistable Circuit Involving SCARECROW-
1123 RETINOBLASTOMA Integrates Cues to Inform Asymmetric Stem Cell
1124 Division, *Cell*. 150 (2012) 1002–1015. doi:10.1016/j.cell.2012.07.017.

1125 [67] L. Di Laurenzio, J. Wysocka-Diller, J.E. Malamy, L. Pysh, Y. Helariutta, G.
1126 Freshour, et al., The SCARECROW Gene Regulates an Asymmetric Cell
1127 Division That Is Essential for Generating the Radial Organization of the
1128 *Arabidopsis* Root, *Cell*. 86 (1996) 423–433. doi:10.1016/S0092-
1129 [8674\(00\)80115-4](https://doi.org/10.1016/S0092-8674(00)80115-4).

1130 [68] Y. Helariutta, H. Fukaki, J. Wysocka-Diller, K. Nakajima, J. Jung, G.
1131 Sena, et al., The SHORT-ROOT Gene Controls Radial Patterning of the
1132 *Arabidopsis* Root through Radial Signaling, *Cell*. 101 (2000) 555–567.
1133 doi:10.1016/S0092-8674(00)80865-X.

1134 [69] Y. Long, W. Smet, A. Cruz-Ramírez, B. Castelijns, W. de Jonge, A.P.
1135 Mähönen, et al., *Arabidopsis* BIRD Zinc Finger Proteins Jointly Stabilize
1136 Tissue Boundaries by Confining the Cell Fate Regulator SHORT-ROOT
1137 and Contributing to Fate Specification, *The Plant Cell*. 27 (2015) 1185–
1138 1199. doi:10.1105/tpc.114.132407.

1139 [70] M.A. Moreno-Risueno, R. Sozzani, G.G. Yardımcı, J.J. Petricka, T.
1140 Vernoux, I. Blilou, et al., Transcriptional control of tissue formation
1141 throughout root development, *Science*. 350 (2015) 426–430.
1142 doi:10.1126/science.aad1171.

1143 [71] T. KOBAYASHI, L. CHEN, K. AIHARA, Modeling Genetic Switches with

- 1144 Positive Feedback Loops, *Journal of Theoretical Biology*. 221 (2003)
 1145 379–399. doi:10.1006/jtbi.2003.3190.
- 1146 [72] G.T.S. Beemster, T.I. Baskin, Analysis of Cell Division and Elongation
 1147 Underlying the Developmental Acceleration of Root Growth in
 1148 *Arabidopsis thaliana*, *Plant Physiology*. 116 (1998) 1515–1526.
 1149 doi:10.1104/pp.116.4.1515.
- 1150 [73] L.M. Blakely, T.A. Evans, Cell dynamics studies on the pericycle of radish
 1151 seedling roots, *Plant Science Letters*. 14 (1979) 79–83.
 1152 doi:10.1016/0304-4211(79)90158-5.
- 1153 [74] J.E. Malamy, P.N. Benfey, Organization and cell differentiation in lateral
 1154 roots of *Arabidopsis thaliana*, *Development*. 124 (1997) 33–44.
- 1155 [75] S. Mangan, U. Alon, Structure and function of the feed-forward loop
 1156 network motif, *Proceedings of the National Academy of Sciences*. 100
 1157 (2003) 1–6. doi:www.pnas.org/cgi/doi/10.1073/pnas.2133841100.
- 1158 [76] K.N. Chang, S. Zhong, M.T. Weirauch, G. Hon, M. Pelizzola, H. Li, et al.,
 1159 Temporal transcriptional response to ethylene gas drives growth hormone
 1160 cross-regulation in *Arabidopsis*, *eLife Sciences*. 2 (2013) e00675.
- 1161 [77] J. Hirsch, E. Marin, M. Floriani, S. Chiarenza, P. Richaud, L. Nussaume,
 1162 et al., Phosphate deficiency promotes modification of iron distribution
 1163 in *Arabidopsis* plants, *Biochimie*. 88 (2006) 1767–1771.
 1164 doi:10.1016/j.biochi.2006.05.007.
- 1165 [78] V. Rai, R. Sanagala, B. Sinilal, S. Yadav, A.K. Sarkar, P.K. Dantu, et al.,
 1166 Iron Availability Affects Phosphate Deficiency-Mediated Responses, and
 1167 Evidence of Cross-Talk with Auxin and Zinc in *Arabidopsis*, *Plant Cell*
 1168 *Physiol*. 56 (2015) 1107–1123. doi:10.1093/pcp/pcv035.
- 1169 [79] Y.S. Kim, M. Lee, J.-H. Lee, H.-J. Lee, C.-M. Park, The unified ICE–CBF
 1170 pathway provides a transcriptional feedback control of freezing tolerance
 1171 during cold acclimation in *Arabidopsis*, *Plant Molecular Biology*. 89 (2015)
 1172 187–201. doi:10.1007/s11103-015-0365-3.
- 1173 [80] J. Balk, S. Lobréaux, Biogenesis of iron–sulfur proteins in plants, *Trends*
 1174 *in Plant Science*. 10 (2005) 324–331. doi:10.1016/j.tplants.2005.05.002.
- 1175 [81] S.A. Sinclair, U. Krämer, The zinc homeostasis network of land plants,
 1176 *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research*. 1823
 1177 (2012) 1553–1567. doi:10.1016/j.bbamcr.2012.05.016.
- 1178 [82] N. Andrés-Colás, A. Perea-García, S. Puig, L. Peñarrubia, Deregulated
 1179 Copper Transport Affects *Arabidopsis* Development Especially in the
 1180 Absence of Environmental Cycles, *Plant Physiology*. 153 (2010) 170–
 1181 184. doi:10.1104/pp.110.153676.
- 1182 [83] P.A. Salomé, M. Oliva, D. Weigel, U. Krämer, Circadian clock adjustment
 1183 to plant iron status depends on chloroplast and phytochrome function,
 1184 *The EMBO Journal*. 32 (2013) 511–523. doi:10.1038/emboj.2012.330.
- 1185 [84] A. Garcia-Molina, S. Xing, P. Huijser, Functional characterisation of
 1186 *Arabidopsis* SPL7 conserved protein domains suggests novel regulatory

1187 mechanisms in the Cu deficiency response, BMC Plant Biology 2014
 1188 14:1. 14 (2014) 1. doi:10.1186/s12870-014-0231-5.
 1189 [85] H. Zhang, X. Zhao, J. Li, H. Cai, X.W. Deng, L. Li, MicroRNA408 Is
 1190 Critical for the HY5-SPL7 Gene Network That Mediates the Coordinated
 1191 Response to Light and Copper, The Plant Cell. 26 (2014) 4933–4953.
 1192 doi:10.1105/tpc.114.127340.
 1193 [86] D. Selote, R. Samira, A. Matthiadis, J.W. Gillikin, T.A. Long, Iron-Binding
 1194 E3 Ligase Mediates Iron Response in Plants by Targeting Basic Helix-
 1195 Loop-Helix Transcription Factors, Plant Physiology. 167 (2015) 273–286.
 1196 [87] X. Li, H. Zhang, Q. Ai, G. Liang, D. Yu, Two bHLH Transcription Factors,
 1197 bHLH34 and bHLH104, Regulate Iron Homeostasis in Arabidopsis
 1198 thaliana, Plant Physiology. 170 (2016) 2478–2493.
 1199 doi:10.1104/pp.15.01827.
 1200 [88] T.A. Long, H. Tsukagoshi, W. Busch, B. Lahner, D.E. Salt, P.N. Benfey,
 1201 The bHLH Transcription Factor POPEYE Regulates Response to Iron
 1202 Deficiency in Arabidopsis Roots, The Plant Cell. 22 (2010) 2219–2236.
 1203 doi:10.1105/tpc.110.074096.
 1204 [89] I.-C. Pan, H.-H. Tsai, Y.-T. Cheng, T.-N. Wen, T.J. Buckhout, W.
 1205 Schmidt, Post-Transcriptional Coordination of the Arabidopsis Iron
 1206 Deficiency Response is Partially Dependent on the E3 Ligases RING
 1207 DOMAIN LIGASE1 (RGLG1) and RING DOMAIN LIGASE2 (RGLG2),
 1208 Mol Cell Proteomics. 14 (2015) 2733–2752.
 1209 doi:10.1074/mcp.M115.048520.
 1210 [90] A.G.L. Assunção, E. Herrero, Y.-F. Lin, B. Huettel, S. Talukdar, C.
 1211 Smaczniak, et al., Arabidopsis thaliana transcription factors bZIP19 and
 1212 bZIP23 regulate the adaptation to zinc deficiency, Proceedings of the
 1213 National Academy of Sciences. 107 (2010) 10296–10301.
 1214 doi:10.1073/pnas.1004788107.
 1215 [91] L. Peñarrubia, N. Andrés-Colás, J. Moreno, S. Puig, Regulation of copper
 1216 transport in Arabidopsis thaliana: a biochemical oscillator? J Biol Inorg
 1217 Chem. 15 (2009) 29–36. doi:10.1007/s00775-009-0591-8.
 1218
 1219

FIGURE LEGENDS

Figure 1: Emergent properties of transcription factor networks (TFNs). Gene regulatory networks include combinations of network motifs (highlighted in color) that influence the transcriptional responses associated with stress and developmental processes. These network motifs engender the complex behaviors seen in these systems such as the emergence of bistability, bifurcations, and sustained oscillations.

Figure 2: Common network motifs. (A) Simple regulation is defined by activation ($\uparrow$) or inhibition ($\perp$) of gene expression. (B) Feed-forward loops occur when the final output is regulated by multiple transcription factors that occur earlier in the transcriptional cascade. (C) Autoregulation occurs when a transcription factor positively ($\uparrow$) or negatively ($\perp$) regulates itself. (D) Negative and Positive Feedback loops. The overall sign of the loop is dictated by the number of inhibitory relationships. An even number indicates an overall positive feedback motif while an odd number indicates an overall negative feedback motif.

Figure 3: Feed-forward loops. (A) Coherence of feed-forward loops is determined by the overall sign of the inputs to the final node. (B) Using Boolean Logic (1 'on'; 0 'off') in feed-forward loops, the output of the final node (reflected in the color of node Z; black (1) or white (0)), is dependent on the sign of the inputs from nodes X and Y (indicated by black (1) or white (0) arrows) and the type of Boolean logic that exists between the two nodes (AND, OR, or Exclusive OR (XOR)).

Figure 4: Root development and abiotic stress regulatory networks contain motifs associated with emergent properties. (A) The regulatory network that influences the asymmetric cell division that gives rise to ground tissue formation contains two Logical AND activations of downstream targets, and a positive feedback loop. (B) The drought regulatory network contains several motifs; two auto activators, an inhibitory relationship that leads to a positive feedback loop, and a coherent type II feed-forward loop. (C) The cold acclimation regulatory network contains an unequal Logical Exclusive OR (XOR) activation and an indirect negative feedback loop. (D-E) The iron deficiency regulatory network contains (D) an indirect coherent type II feed-forward loop with a Logical AND activation of downstream targets, and (E) an indirect inhibitory relationship, coupled with a direct inhibitory relationship which leads to a positive feedback motif. A coherent type II feed-forward loop, and an autoregulation (inhibitory or activating effect unknown) motif. In systems theory, each of these motifs are associated with emergent behavior and in some cases have been verified in plant biological systems. (F) Coherent type II feed-forward motifs are associated with delayed response, auto activation is associated with bistability, negative feedback is associated with returning transcript levels to their original steady state, and

1263 double inhibitory feedback loops, which lead to a positive feedback summation, is
1264 associated with bistability.
1265

1266
1267

1268
1269

1270
1271

1272