Systems Biology Update: Cell Type–Specific Transcriptional Regulatory Networks

Article *in* Plant physiology · December 2009
DOI: 10.1104/pp.109.148668 · Source: PubMed

2 authors:

Li Pu
University of California, Berkeley

Siobhan M Brady
University of California, Davis

Systems Biology Update: Cell Type-Specific Transcriptional Regulatory Networks

Li Pu and Siobhan Brady*

Department of Plant Biology and Genome Center, University of California, Davis, California 95616

Plant cells use regulatory networks composed of numerous components, such as DNA, RNA, proteins, and small molecules, to regulate multiple biological processes, allowing plants to adapt to changing environments or to respond to developmental cues. The availability of high-throughput experimental methods enables researchers to determine the expression levels for thousands of genes and protein-protein or protein-DNA interactions. Systems biology approaches can allow scientists to integrate these large amounts of information and to understand the properties of these biological systems in specific cells or tissues. Dynamic mathematical modeling approaches used to characterize plant transcriptional networks can reveal emergent properties of these networks. This review highlights some currently available methodologies used to obtain systems-scale data such as laser capture microdissection (LCM), fluorescence-activated cell sorting (FACS), chromatin immunoprecipitation (ChIP)-on-chip, proteomics, and modeling approaches that are most useful to explore plant transcriptional networks at the cellular level. We also provide two examples of transcriptional networks in single cell types and detail how such methods and data sets have been used to map and reveal emergent properties of gene regulatory networks that regulate cell identity specification.

In the 21st century, our species and planet are facing many urgent problems—including our diminishing supply of nonrenewable fuel sources, as well as food and water shortages. However, plant biology has the potential to provide solutions to these global crises. To address these problems, efforts to understand and model how a plant responds to environmental stimuli or genetic manipulation using systems biology approaches were proposed (Raikhel and Coruzzi, 2003). Genome-scale acquisition of data describing gene and protein expression and interactions has been and continues to be collected. To better understand these and other large-scale data sets, systems biology approaches were utilized. Systems biology aims to discover emergent properties that arise from the analysis of interactions between components in a biological system. These properties are revealed by integrating, modeling, and analyzing the interactions between all components using network theory, systematic perturbation, and monitoring of the modeled components, and model refinement to reconcile discrepancies between the experimental observations and the existing model (Ideker et al., 2001; Long et al., 2008). Ideally these approaches are best utilized with the full complement of components that exist within gene regulatory networks. However, systems approaches have also been utilized with small cell type-specific regulatory networks elucidated using molecular genetic approaches. The identity of individual plant cells are specified by and respond to changing environments by connecting numerous components, such as DNA, RNA, proteins, and small molecules, within regulatory networks to coordinate multiple biological functions (Brady et al., 2007; Dinneny et al., 2008).

What types of -omics data are ideally needed to reconstruct plant transcriptional regulatory networks at the cellular level? Transcriptomes are obtained by measuring whole-genome expression using microarrays or short-read sequencing methods (Brady et al., 2006; Iyer-Pascuzzi et al., 2009). Proteomes are obtained by two-dimensional gel electrophoresis and mass spectrometry (Raikhel and Coruzzi, 2003; Baginsky and Gruissem, 2006). Regulatory elements are regions of DNA or RNA where regulators such as transcription factors (TFs) bind preferentially to regulate the expression of target genes. These elements are often identified using computational methods coupled with gene expression data (Long et al., 2008; Priest et al., 2009). The metabolome represents the collection of all metabolites in a biological organism, which are additionally the end products of transcriptional regulation, and which can influence transcription as signaling molecules (Brady et al., 2006; Schauer and Fernie, 2006; Long et al., 2008). Additional signaling molecules like carbohydrates and lipids can be collected, analyzed, and processed at the systems level (Watson, 2006; Aoki-Kinoshita, 2008). Recent reviews have discussed techniques used to identify components for use in plant systems biology such as genome sequencing, microarrays, pull-down assays, the yeast one-hybrid system, and other important tools (Brady et al., 2006; Albert, 2007; Long et al., 2008).

A remaining challenge for plant biologists is to integrate and understand the biological properties that emerge from the interactions of cellular components. Recently, plant biologists have been able to discover components of transcriptional networks in individual cell types using systems biology approaches. These have been highlighted in Table I. Some prominent reviews

* Corresponding author; e-mail sbrady@ucdavis.edu.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantphysiol.org) is: Siobhan Brady (sbrady@ucdavis.edu).

www.plantphysiol.org/cgi/doi/10.1104/pp.109.148668

Table I. *Characterization of transcriptional regulatory networks at the cellular level*
Asterisk (*) indicates studies of great interest.

Method	Tissue/Cell Type	Plant Species	Reference
ChIP-chip	Trichome	Arabidopsis	Morohashi and Grotewold (2009)*
ChIP-qPCR	Trichome	Arabidopsis	Wang et al. (2007)
ChIP-qPCR	Root endodermis	Arabidopsis	Levesque et al. (2006), Cui et al. (2007)*
Microarray and FACS	Root (five tissues)	Arabidopsis	Dinneny et al. (2008)*
Microarray and FACS	Root (five tissues)	Arabidopsis	Gifford et al. (2008)*
Microarray and FACS	SAM	Arabidopsis	Yadav et al. (2009)
Microarray and FACS	Sperm	Arabidopsis	Borges et al. (2008)
Microarray and LCM	Root (14 cell types)	Arabidopsis	Brady et al. (2007)*
Microarray and LCM	Endosperm	Arabidopsis	Day et al. (2008)
Microarray and LCM	Embryo	Arabidopsis	Spencer et al. (2007)
Microarray and LCM	Root, shoot, and embryo (40 cell types)	Rice	Jiao et al. (2009)*
Microarray and LCM	SAM	Maize	Ohtsu et al. (2007), Zhang et al. (2007), Brooks et al. (2009)*
Microarray and LCM	Fiber cell	Cotton	Wu et al. (2007)
Microarray and modeling approaches	Pollen tube	Arabidopsis	Qin et al. (2009)*
Microarray, LCM, and proteomics	Pericycle cells	Maize	Dembinsky et al. (2007)*
Modeling approaches	Root hair	Arabidopsis	Savage et al. (2008)*
Modeling approaches	Leaf and root epidermis	Arabidopsis	Benitez et al. (2007, 2008)
Modeling approaches	Root	Arabidopsis	Cartwright et al. (2009)
Proteomics	Root, leaves, and flower	Arabidopsis	Baerenfaller et al. (2008)*
Proteomics	Guard cell	Arabidopsis	Zhao et al. (2008b)
Proteomics	Pollen	Rice	Dai et al. (2007)
Proteomics	Root hair	Soybean	Brechenmacher et al. (2009)

have additionally been published regarding one or two aspects of systems biology approaches (Alon, 2007; Alvarez-Buylla et al., 2007; Camacho et al., 2007; Busser et al., 2008). Here, we will describe some currently well-used systems approaches to explore plant transcriptional networks at the cellular level. We will also describe two examples that demonstrate how these methods have been used to characterize gene regulatory networks that regulate cell identity specification, and that reveal emergent properties of these networks.

GENOME-SCALE METHODS USED TO IDENTIFY COMPONENTS OF CELL-TYPE TRANSCRIPTIONAL REGULATORY NETWORKS

Plant growth and development depends, to a large degree, on the tissue- or cell-type-specific expression of genes. Although the use of large-scale microarray technology has led to the generation of tremendous amounts of plants gene expression data in whole plants or organs, this data is unlikely to provide information about gene expression in specialized tissue or cell types (Kilian et al., 2007; Goda et al., 2008). More recently, two techniques in conjunction with microarrays were developed to study transcriptional changes in individual cell types and tissues within the plant.

LCM

LCM is a powerful tool allowing the rapid and precise isolation of specific populations of cells or even individual cells from a heterogeneous tissue based on established histological identification. LCM combined with DNA microarray analysis was used to identify a high-resolution expression map of the syncytial stage of Arabidopsis (*Arabidopsis thaliana*) endosperm development at 4 d after pollination (Day et al., 2008). Using this technique, Jiao et al. obtained a cell-type transcriptome atlas that includes 40 cell types from rice (*Oryza sativa*) shoot, root, and the germinating seed at several developmental stages (Jiao et al., 2009). This method has been used to isolate cell-type populations in maize (*Zea mays*) and cotton (*Gossypium hirsutum*; Dembinsky et al., 2007; Ohtsu et al., 2007; Wu et al., 2007; Zhang et al., 2007; Brooks et al., 2009). Cells isolated by this technology have also been used to characterize the cell-type-specific proteome in the maize root. This study defined the distinct molecular events during the specification of cell-cycle-competent pericycle cells prior to their first division, and also demonstrates that pericycle specification and lateral root initiation might be controlled by a different set of genes (Dembinsky et al., 2007).

FACS

Researchers have used FACS to isolate RNA from cell-type-specific GFP-marked root populations (Birnbaum et al., 2003, 2005). Using this technique, high spatial resolution gene expression data have been obtained for nearly all cell types in the root. This analysis revealed dominant expression patterns between unrelated cell types and was used to infer novel cellular function (Brady et al., 2007). This method has also been used to examine the response to high salt and iron deficiency in specific cell types and developmental stages of the root (Dinneny et al., 2008). Most responses to salt and iron were cell-type specific and were dependent on environmental conditions. This study revealed that the level of conservation among biological functions enriched in cell types under high salt and low iron conditions varied according to cell type and stress (Dinneny et al., 2008). A cell-type-specific transcriptional network underlying nitrogen responses in the root has also been revealed using FACS that mediates lateral root outgrowth (Gifford et al., 2008).

ChIP

DNA-binding proteins perform a variety of important functions in cells, including transcriptional regulation, chromosome maintenance, replication, and DNA repair. The interactions between TFs and their DNA-binding sites are an integral part of transcriptional regulatory networks. ChIP is a well-established procedure used to investigate interactions between proteins and DNA (Buck and Lieb, 2004). More importantly, ChIP assays in combination with other methods have been used to determine the interaction between TFs and DNA from a few potential individual targets to genome-wide surveys. Coupled with whole-genome DNA microarrays, ChIP-on-chip (also called ChIP-chip) can provide a whole-genome view of DNA-binding sites for any given protein that occurs in vivo. This powerful procedure has been adapted to identify a set of about 20 genes that are targets of GLABRA3 (GL3) and GL1 in the Arabidopsis trichome (Morohashi and Grotewold, 2009). An additional and less comprehensive method utilizes ChIP in conjunction with quantitative real-time PCR (ChIP-qPCR). This technique combined with FACS has been used to characterize the targets of SHORTROOT (SHR) and SCARECROW TFs that specify the endodermis in the Arabidopsis root (Levesque et al., 2006; Cui et al., 2007). ChIP-qPCR was also used to demonstrate the interaction of TRICHOMELESS1 and the cis-acting regulatory elements of *GL1* that regulate trichome cell specification in Arabidopsis that provided, to our knowledge, the first molecular and genetic evidence that an R3 MYB may negatively regulate trichome cell specification in a novel manner by directly suppressing the transcription of *GL1* (Wang et al., 2007).

PROTEOMICS

Every cell within the plant contains the same DNA, yet different cells appear committed to specialized tasks. How is this possible? The answer lies in the differential regulation of gene expression and protein synthesis. This process begins with transcription and is completed with protein translation and subsequent posttranslational modification. Thus, proteomic approaches are key to investigating transcriptional networks at the cellular level. Improved high-throughput proteomics techniques has shifted attention to protein profiling, which attempts to identify all proteins that are present in a particular tissue or cell (Baginsky and Gruissem, 2006). Recently, an Arabidopsis proteome map has been established in different organs, developmental stages, and in undifferentiated cultured cells (Baerenfaller et al., 2008). Proteins in gene ontology categories for intracellular protein transport, response to oxidative stress, and toxin catabolic process were represented in their analyses (Baerenfaller et al., 2008). An additional dataset that profiles the proteome of four organs in Arabidopsis also sampled the phosphorylated proteins (Castellana et al., 2008). These databases provide information about genome activity and proteome assembly for plant systems biology approaches. In Arabidopsis, cell type-specific proteome maps have only been established in guard cells. Proteomic methods were used to identify 1,734 unique Arabidopsis guard cell proteins including 336 proteins not previously represented in transcriptome analyses of guard cells and 52 proteins classified as signaling proteins by gene ontology analysis (Zhao et al., 2008b). Two-dimensional PAGE coupled with mass spectrometry has been used to identify proteins enriched in a single root hair cell of soybean (*Glycine max*). The proteins identified are involved not only in basic cell metabolism but also in functions more specific to the single root hair cell, including water and nutrient uptake, vesicle trafficking, and hormone and secondary metabolism that provided insight into the metabolic activities of a single, differentiated plant cell type (Brechenmacher et al., 2009).

META-ANALYSES

High-throughput microarray technology has become a popular tool for large-scale gene expression analysis. As a result, there are rapidly growing collections of available data sets that can be used for subsequent analysis. Meta-analysis consists of a set of statistical techniques that have been used to combine results from several independent microarray experiments. This technique appears to be a practical solution to maximize the use of data available from each experiment. This method has led to the identification of cell-type transcriptional networks (Hong and Breitling, 2008; Iyer-Pascuzzi et al., 2009). For example, eight direct targets of SHR were identified by combining three different microarray expression datasets. Four of eight putative targets were further confirmed by their binding of SHR by ChIP-qPCR in vivo (Levesque et al., 2006). SHR controls radial patterning of the root by regulating endodermis and stem cell niche specification. SHR is also expressed in the vasculature of the root, and this meta-analysis identified targets of SHR in the vasculature and a new function of SHR in vasculature development. Furthermore, meta-analysis has been used to establish a minimal set of genes uniquely associated with the formation of trichomes controlled by GL1 and/or GL3 in Arabidopsis (Morohashi and Grotewold, 2009). Therefore, meta-analysis provides researchers with a powerful tool to interrogate existing databases for further exploration of transcriptional regulatory networks in specific tissues or cell types.

MODELING APPROACHES

Mathematical and computational tools can be used on data from multiple sources (including genome-scale data and small-scale data) to dynamically model gene regulatory networks (Alvarez-Buylla et al., 2007; Busser et al., 2008). Network theory can then be used to analyze these models to better understand complex systems that have multiple elements and interactions. These predictive frameworks are an extremely important aspect of systems biology that can greatly improve our understanding of the local dynamics of cell-type-specific transcriptional circuits. Boolean and Bayesian network (BN) models are two commonly used modeling approaches to describe dynamic regulatory networks. A Boolean network is a system of binary-state nodes (representing genes) with edges between each node that represent regulatory interactions. One of two states is assigned (on/off) that represent the status of a gene being active or inactive, respectively. The change in state of each regulated node is generally described by a logical function using the Boolean operators and, or, and not. Boolean models can predict dynamic trends within the system at a particular time point (Albert, 2007; Price and Shmulevich, 2007). A BN is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph (Albert, 2007; Price and Shmulevich, 2007). A dynamic BN extends the notion of a BN to model the stochastic evolution of a set of random variables over time (Camacho et al., 2007). A probabilistic Boolean modeling approach has recently been used to analyze the gene regulatory network that underlies Arabidopsis root epidermis patterning (Benitez et al., 2008; Savage et al., 2008). The ability of these modeling approaches to reveal emergent properties of this gene regulatory network in determining cell identity specification will be further discussed.

ROOT HAIR PATTERNING IN ARABIDOPSIS

The appropriate specification and patterning of cell types are fundamental features of plant development. Plant root hairs form an important surface that absorbs water and nutrients, and protects the root from water loss and insect and pathogen invasion. In Arabidopsis, root hair cells are specified in a position-dependent manner. All epidermal cells located above two underlying cortical cells (designated the H position) develop as hair cells and cells located above a single cortical cell (designated the N position) adopt the non-hair fate. The formation of root hair cells and non-hair cells in the epidermis of plant roots provides an excellent model to study cell-type patterning because root epidermal cells are accessible, their developmental process can be accurately analyzed, and they differentiate in a predictable gradient along the root axis (Schiefelbein et al., 2009). While some genome-scale expression data has been collected for root hairs (Brady et al., 2007), primarily small-scale molecular genetics studies have resulted in a well-characterized network regulating hair and non-hair cell specification in the Arabidopsis root. Modeling approaches, however, have addressed the hypothesis that this network specifies cell identity via a mutual support mechanism with *WEREWOLF* (*WER*) autoregulation, and lateral movement of CAPRICE (CPC) and GL3 (Fig. 1B; Savage et al., 2008).

Factors that act in root hair cell patterning in Arabidopsis include an R2R3 MYB-type TF, WER; a WD-repeat protein, TRANSPARENT TESTA GLABRA1 (TTG1); two partially redundant basic helix-loop-helix (bHLH) proteins, GL3 and ENHANCER OF GLABRA3 (EGL3); a homeodomain protein, GL2; and three MYB proteins, CPC, TRIPTYCHON (TRY), and ENHANCER OF TRIPTYCHON AND CAPRICE1 (ETC1). It has been proposed that WER, GL3/EGL3, and TTG1 form a transcriptional complex involved in the promotion of non-hair cell specification (Fig. 1, B and C). This regulatory complex is preferentially expressed in the non-hair cell position, and functions to induce the expression of *GL2* that promotes the non-hair fate, while the same complex simultaneously activates the expression of *CPC*, *TRY*, and *ETC1* in N cells (Ishida et al., 2008; Schiefelbein et al., 2009). CPC, TRY, and ETC1 then move into neighboring epidermal cells where they compete with WER for physical interaction with GL3/EGL3 to repress *WER* and *GL2* expression, resulting in H-cell specification (Tominaga et al., 2007; Ishida et al., 2008). Positional determination of N- or H-cell identity acquisition is also influenced by SCRAMBLED (SCM), a Leu-rich receptor-like kinase that is proposed to interpret a positional signal from the underlying cortical cells (Fig. 1, B and C; Kwak and Schiefelbein, 2008; Schiefelbein et al., 2009). Autoregulation of *WER* expression in N cells has also been proposed to be necessary for epidermal patterning within this network (Fig. 1C).

Two alternative forms of WER regulation were considered in the modeling of this network using Boolean formalism: (1) local WER self activation in N cells implemented by enhancement of *WER* transcription by the WER-EGL3/GL3-TTG1 complex or (2) *WER* transcription activated uniformly in all epidermal cells (Fig. 1C). In this Boolean modeling approach, components were either expressed or not. For exam-

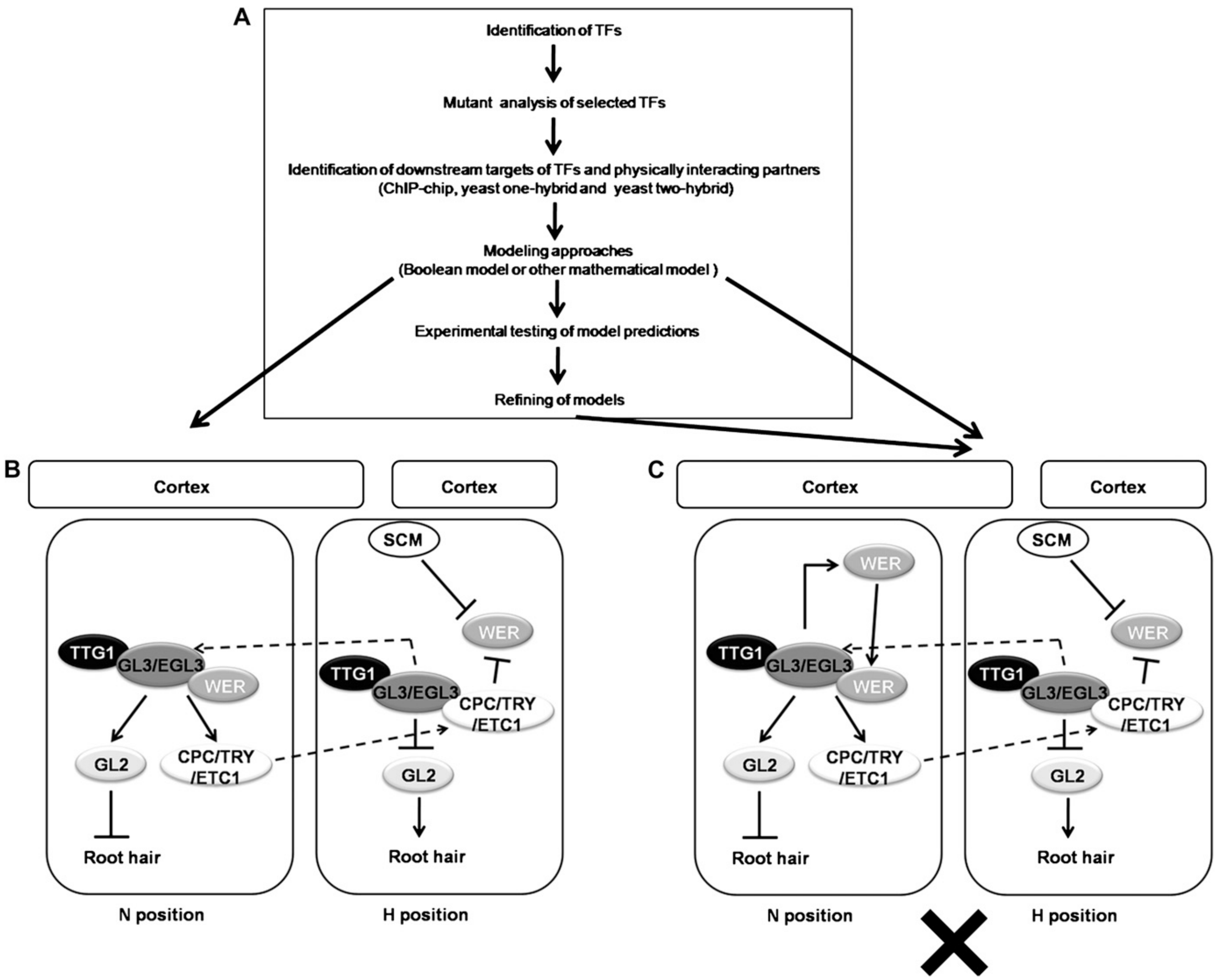

Figure 1. System approaches and two models describing the regulatory network for cell-type patterning in the Arabidopsis root epidermis. A, System approaches used to study the transcriptional networks in root hairs. TFs involved in root hair development are first identified. Mutants of cell-type-specific or tissue-specific TFs are analyzed and transcriptionally profiled to detect potential downstream targets of the selected TFs. The physical interactions between the TFs and their potential targets (TFs or promoters) determined by in vitro tests such as yeast one-hybrid assays, yeast two-hybrid assays, or in vivo by ChIP-chip analysis. Subsequently, TFs and targets are integrated by modeling approaches such as Boolean modeling. This model should then be tested experimentally to determine its predictions to recapitulate experimental observations. B and C, Two models characterizing regulatory networks of cell-type patterning in the Arabidopsis root epidermis—a mutual support model (B) and autoactivation of WER model (C). In model of mutual support, accumulation of the WER-GL3/EGL3-TTG transcriptional complex in N cells promotes *GL2* that directs non-hair cell differentiation. This complex also activates *CPC/TRY/ETC1* that move into neighboring cells where they compete with WER for binding to GL3/EGL3 that inhibit GL2 expression and result in specification of hair cell fate. In H cells, GL3/EGL3 move to N cells where they form part of the WER complex that induces expression of *CPC/TRY/ETC1* and *GL2*. The SCM receptor preferentially represses WER transcription in the H-cell position to influence the pattern formation (B). In the model of WER autoregulation, WER activates its expression in N cells implemented by enhancement of *WER* transcription by the WER-EGL3/GL3-TTG1 complex (C). But this model was rejected after simulations (showed by x). Arrows indicate positive regulation, blunted lines indicate negative regulation, and broken lines indicate intercellular protein movement. After model simulations, and experimental testing, the WER autoregulation modeled network was rejected.

ple, components with only a single positive input were expressed if their direct positive regulator is expressed. In cases where multiple positive and negative inputs exist, these components were given a time-evolving probability of expression that is correlated with their expression abundance in the cell. A positional bias was implemented by expressing SCM only in H cells, resulting in a lower transcription rate of *WER* than in the N position. Using a simulated ring of epidermal cells, these models were able to recapitulate wild-type expression patterns within the epidermis. These models were then perturbed according to experimentally characterized mutations in components of the network. In a simulated *cpc* mutant, however, these two models yielded different expression patterns, with only the mutual support framework, yielding expression patterns similar to experimental observations. To then experimentally demonstrate if WER is able to regulate its own expression, the ability of WER to drive its own expression was characterized in wild type and in a *wer* mutant. Identical spatial expression patterns were observed, suggesting that WER does not autoregulate its expression. To then determine if *WER* is expressed uniformly in all epidermal cells early in development, this same reporter was analyzed in wild type and found to be expressed uniformly within the root meristem. These results demonstrate that *WER* is initially expressed uniformly and its expression is restricted due to *SCM* expression and CPC movement (Fig. 1C). The autoactivation model network was also unable to recapitulate experimentally determined WER expression patterns in a *gl3/egl3* and in a *ttg* mutant. An additional dynamic gene regulatory modeling approach was used to model Arabidopsis root epidermal patterning where each gene can have one of three expression states (0 = off, 1 = mild expression, 2 = strong expression; Benitez et al., 2007, 2008). This model requires *WER* autoregulation and was not able to recapitulate the observed *WER* expression pattern in a *scm* mutant. The authors further provide an alternative regulatory mechanism to determine cell patterning in the root epidermis via the combined activity of a self-activating loop and another input activating *WER*. This mechanism, the authors suggest, allows both models to be plausible and recapitulate experimental data.

Patterning root hairs in Arabidopsis have provided us a comprehensive understanding of the transcriptional network to guide the fate of epidermal cells. However, future research will be necessary to address several unclear issues in this regulatory system. For example, the exact regulatory mechanism for intercellular movement of the CPC protein and GL3/EGL3 proteins is still a mystery. Moreover, the full complement of targets of these TFs in root hairs remains to be determined. Studies that incorporate more sophisticated computational or mathematical models that bridge the dynamics of pattern formation and that are informed by further experimentation will further elaborate the gene regulatory network in root epidermal cell patterning.

TRICHOME PATTERNING IN ARABIDOPSIS

Trichomes of Arabidopsis are developmentally important because they are involved in temperature control, water regulation, and protection against insect herbivores and UV irradiation. Arabidopsis trichomes are single-celled, branched hair-like structures that differentiate from epidermal cells of leaves, stems, and sepals. In contrast to root hair patterning, which is predictable with respect to the underlying cortex and position-dependent cell fate determination, regular trichome patterning is an example of de novo pattern formation (Ishida et al., 2008). The establishment of a single-celled trichome on the leaf also serves as an excellent model system to study the gene regulatory networks involved in cell fate determination. Systems approaches have been used to elucidate the key regulatory molecular interactions that determine trichome spacing.

Similar components of the gene regulatory network that regulate root hair specification also operate during trichome determination (Zhao et al., 2008a; Schiefelbein et al., 2009). They include an R2R3 MYB TF GL1, the bHLH factors GL3 and EGL3, and the WD40-repeat protein TTG1 that form a trichome-promoting complex GL1-GL3/EGL3-TTG1 (Fig. 2). Two downstream targets of this complex that promote trichome formation are the homeobox TF GL2 and the WRKY regulator, TTG2. The negative regulators *TRY*, *CPC*, and *ETC1* are initially activated by this activating complex and ETC1 had been demonstrated to move into neighboring cells while CPC and TRY had been hypothesized to also act cell nonautonomously. These inhibitors can compete with GL1 for binding to the GL1-GL3/EGL3 active complex to suppress trichome initiation and differentiation into hairless cells (Fig. 2; Ishida et al., 2008; Zhao et al., 2008a). A modeling study using ordinary differential equations has eluci-

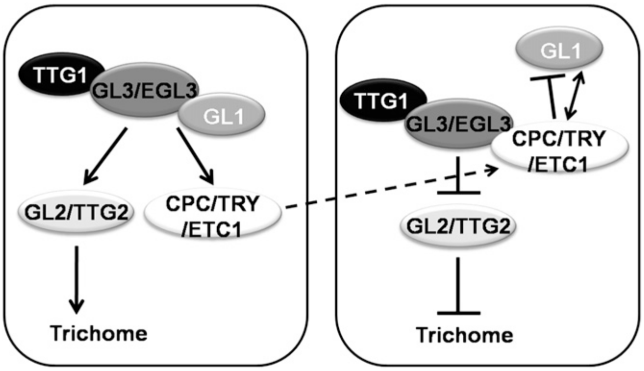

Figure 2. A transcriptional network of cell-type patterning in the Arabidopsis trichome. In trichome cells, the activating complex TTG1-GL3/EGL3-GL1 activates the expression of trichome activators (*GL2/TTG2*) and single MYB inhibitors in the epidermal cell chosen to be a trichome. The inhibitors CPC/TRY/ETC1 then move into the adjacent cells and through physical interactions with GL3/EGL3 competes for GL1 complex formation forming an inactivating complex. This inactivating complex disrupts the function of the activating complex, thereby decreasing the expression of *GL2/TTG2*, and the trichome cell fate is not triggered. TRY also binds to free GL1, thereby preventing the interaction between GL1 and GL3/EGL3. Arrows indicate positive regulation, blunted lines indicate negative regulation, broken arrows indicate intercellular movement of protein, and double arrow indicates interaction.

dated the molecular mechanism by which TRY can exert its inhibitory function (Digiuni et al., 2008). The authors determined experimentally that TRY, in addition to interacting physically with GL3, can also physically interact with GL1. Three alternative models were proposed—a single competitive inhibition model in which TRY binds to free GL3 and blocks the formation of the GL1-GL3/EGL3-TTG1 active complex, a double competitive inhibition model that additionally includes the binding of TRY to free GL1, and an uncompetitive inhibition model where TRY binds only to the activator complex and renders it inactive (Fig. 2). These models were run and a series of GL3 and TRY overexpression experiments performed. The resulting experimental data best matched the single competitive inhibition model simulation data and therefore predicts that the interaction between TRY and GL3 is the most relevant for trichome patterning.

To further determine the number of targets of the active complex, a recent study identified 20 direct targets of both GL1 and GL3 by using ChIP-chip and genome-wide expression analyses (Morohashi and Grotewold, 2009). Among the GL3/GL1 direct targets, at least four TFs, *GL2*, *TTG2*, *SCL8*, and *MYC1* constitute main transcriptional regulators of the network hierarchical structure for trichome differentiation. Temporal induction analyses further determined that many of these genes are expressed early upon GL1 or GL3 induction, suggesting that they play important roles in the early events in differentiation of protodermal cells. For example, *SCL8* encodes a GRAS family TF and *SCL8* mRNA levels peak sharply within the first few hours of GL3 or GL1 induction, to then level off at quantities similar as found in the absence of the regulators, suggesting a need for *SCL8* function at early stages during trichome initiation. *MYC1* encodes a bHLH factor closely related to GL3 and EGL3 and *MYC1* mRNA also accumulates in the early of trichome initiation. Although SCL8 and MYC1 are implicated to be required for trichome development, their functions remain uncharacterized. Additionally, these methods predict novel regulators of trichome formation—a minimal set of 513 genes were identified to be associated with trichome formation and are downstream targets of this transcriptional regulatory cascade (Morohashi and Grotewold, 2009).

PERSPECTIVE

The advent of these large-scale, high-resolution datasets and computational tools provide the means for a better understanding of transcriptional regulatory networks in a cell type or tissue. Methods used to obtain single-cell or cell-type-resolution transcription profiles have indeed become robust enough to facilitate their use. Despite these achievements, however, we still lack sufficient cell-type-specific data needed to completely elucidate these networks including the cell-type-specific small RNA compendium, epigenome, proteome, phosphoproteome, metabolome, and lipidome (Zhang et al., 2006; Henderson and Jacobsen, 2007; Kasschau et al., 2007; Zilberman et al., 2007; Brodersen et al., 2008; Castellana et al., 2008; Cokus et al., 2008; Lister et al., 2008; Zhu, 2008; Popescu et al., 2009). The LCM and cell-sorting approaches have proven efficient to isolate cell-type-specific transcriptomes often via amplification of RNA in up to two cDNA synthesis steps. Unfortunately, however, it is not possible to amplify the proteome, metabolome, or lipidome. Therefore, a limitation exists in obtaining the massive amounts of material needed to generate these types of cell-type-specific data. This limitation is not insurmountable, however, it just requires sufficient plant, financial, and person resources.

In addition, to determine the regulatory logic that underlies these transcriptional regulatory networks in individual cells or cell types, we need to identify the targets of all TFs expressed within a cell type, and their preferential binding sites. The majority of experimental approaches used to elucidate downstream targets of TFs in plants utilize a TF-centered approach, where targets of a single TF of interest are characterized. The targets of all TFs expressed in a cell type need to be determined in an unbiased manner to completely and fully understand the function of these transcriptional networks within the cell. However, to collect this data, tagged versions of all TFs expressed in a cell type ideally under their native promoter would need to be synthesized. An alternate gene-centered approach could be utilized with high-throughput yeast one-hybrid approaches utilizing whole cell-type-specific promoters as bait, and prey TF libraries comprised of TFs expressed in the same cell type (Deplancke et al., 2006; Pruneda-Paz et al., 2009). Studies of this magnitude have been carried out in many other organisms, and have revealed emergent properties of these regulatory networks (Deplancke et al., 2006; Vermeirssen et al., 2007; Wu et al., 2007; Brooks et al., 2009; Jiao et al., 2009). Finally, studies characterizing the nature of these transcriptional interactions (activating or repressing), and the temporal dynamics of these interactions are needed to correctly understand flux within these regulatory networks.

Bioinformatic tools also need to be further developed that enable the user to integrate and visualize these diverse data in the appropriate statistical manner. These will enable the construction of integrated models that provide a predictive framework to better understand plant cell type or tissue development. A further challenge for a plant system biologist is in the generation of modeling tools that allows one to build comprehensive models at the appropriate cellular and temporal scale, such as those currently being developed by the iPlant collaborative. We are certainly at the brink of collecting sufficient transcriptomic data to begin to map the transcriptional regulatory networks that regulate the development and function of plant cells and tissues. However, considerably more data is needed to obtain a comprehensive view of these networks. Once obtained, modeling of these data will enable the improvement of crop yield, biofuel production, and other urgent needs.

Received October 1, 2009; accepted November 30, 2009; published December 4, 2009.

LITERATURE CITED

Albert R (2007) Network inference, analysis, and modeling in systems biology. Plant Cell **19:** 3327–3338

Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet **8:** 450–461

Alvarez-Buylla ER, Benitez M, Davila EB, Chaos A, Espinosa-Soto C, Padilla-Longoria P (2007) Gene regulatory network models for plant development. Curr Opin Plant Biol **10:** 83–91

Aoki-Kinoshita KF (2008) An introduction to bioinformatics for glycomics research. PLoS Comput Biol **4:** e1000075

Baerenfaller K, Grossmann J, Grobei MA, Hull R, Hirsch-Hoffmann M, Yalovsky S, Zimmermann P, Grossniklaus U, Gruissem W, Baginsky S (2008) Genome-scale proteomics reveals Arabidopsis thaliana gene models and proteome dynamics. Science **320:** 938–941

Baginsky S, Gruissem W (2006) Arabidopsis thaliana proteomics: from proteome to genome. J Exp Bot **57:** 1485–1491

Benitez M, Espinosa-Soto C, Padilla-Longoria P, Alvarez-Buylla ER (2008) Interlinked nonlinear subnetworks underlie the formation of robust cellular patterns in Arabidopsis epidermis: a dynamic spatial model. BMC Syst Biol **2:** 98

Benitez M, Espinosa-Soto C, Padilla-Longoria P, Diaz J, Alvarez-Buylla ER (2007) Equivalent genetic regulatory networks in different contexts recover contrasting spatial cell patterns that resemble those in Arabidopsis root and leaf epidermis: a dynamic model. Int J Dev Biol **51:** 139–155

Birnbaum K, Jung JW, Wang JY, Lambert GM, Hirst JA, Galbraith DW, Benfey PN (2005) Cell type-specific expression profiling in plants via cell sorting of protoplasts from fluorescent reporter lines. Nat Methods **2:** 615–619

Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN (2003) A gene expression map of the Arabidopsis root. Science **302:** 1956–1960

Borges F, Gomes G, Gardner R, Moreno N, McCormick S, Feijo JA, Becker JD (2008) Comparative transcriptomics of Arabidopsis sperm cells. Plant Physiol **148:** 1168–1181

Brady SM, Long TA, Benfey PN (2006) Unraveling the dynamic transcriptome. Plant Cell **18:** 2101–2111

Brady SM, Orlando DA, Lee JY, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN (2007) A high-resolution root spatiotemporal map reveals dominant expression patterns. Science **318:** 801–806

Brechenmacher L, Lee J, Sachdev S, Song Z, Nguyen TH, Joshi T, Oehrle N, Libault M, Mooney B, Xu D, et al (2009) Establishment of a protein reference map for soybean root hair cells. Plant Physiol **149:** 670–682

Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science **320:** 1185–1190

Brooks L III, Strable J, Zhang X, Ohtsu K, Zhou R, Sarkar A, Hargreaves S, Elshire RJ, Eudy D, Pawlowska T, et al (2009) Microdissection of shoot meristem functional domains. PLoS Genet **5:** e1000476

Buck MJ, Lieb JD (2004) ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. Genomics **83:** 349–360

Busser BW, Bulyk ML, Michelson AM (2008) Toward a systems-level understanding of developmental regulatory networks. Curr Opin Genet Dev **18:** 521–529

Camacho D, Vera Licona P, Mendes P, Laubenbacher R (2007) Comparison of reverse-engineering methods using an in silico network. Ann N Y Acad Sci **1115:** 73–89

Cartwright DA, Brady SM, Orlando DA, Sturmfels B, Benfey PN (2009) Reconstructing spatiotemporal gene expression data from partial observations. Bioinformatics **25:** 2581–2587

Castellana NE, Payne SH, Shen Z, Stanke M, Bafna V, Briggs SP (2008) Discovery and revision of Arabidopsis genes by proteogenomics. Proc Natl Acad Sci USA **105:** 21034–21038

Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature **452:** 215–219

Cui H, Levesque MP, Vernoux T, Jung JW, Paquette AJ, Gallagher KL, Wang JY, Blilou I, Scheres B, Benfey PN (2007) An evolutionarily conserved mechanism delimiting SHR movement defines a single layer of endodermis in plants. Science **316:** 421–425

Dai S, Chen T, Chong K, Xue Y, Liu S, Wang T (2007) Proteomics identification of differentially expressed proteins associated with pollen germination and tube growth reveals characteristics of germinated Oryza sativa pollen. Mol Cell Proteomics **6:** 207–230

Day RC, Herridge RP, Ambrose BA, Macknight RC (2008) Transcriptome analysis of proliferating Arabidopsis endosperm reveals biological implications for the control of syncytial division, cytokinin signaling, and gene expression regulation. Plant Physiol **148:** 1964–1984

Dembinsky D, Woll K, Saleem M, Liu Y, Fu Y, Borsuk LA, Lamkemeyer T, Fladerer C, Madlung J, Barbazuk B, et al (2007) Transcriptomic and proteomic analyses of pericycle cells of the maize primary root. Plant Physiol **145:** 575–588

Deplancke B, Mukhopadhyay A, Ao W, Elewa AM, Grove CA, Martinez NJ, Sequerra R, Doucette-Stamm L, Reece-Hoyes JS, Hope IA, et al (2006) A gene-centered C. elegans protein-DNA interaction network. Cell **125:** 1193–1205

Digiuni S, Schellmann S, Geier F, Greese B, Pesch M, Wester K, Dartan B, Mach V, Srinivas BP, Timmer J, et al (2008) A competitive complex formation mechanism underlies trichome patterning on Arabidopsis leaves. Mol Syst Biol **4:** 217

Dinneny JR, Long TA, Wang JY, Jung JW, Mace D, Pointer S, Barron C, Brady SM, Schiefelbein J, Benfey PN (2008) Cell identity mediates the response of Arabidopsis roots to abiotic stress. Science **320:** 942–945

Gifford ML, Dean A, Gutierrez RA, Coruzzi GM, Birnbaum KD (2008) Cell-specific nitrogen responses mediate developmental plasticity. Proc Natl Acad Sci USA **105:** 803–808

Goda H, Sasaki E, Akiyama K, Maruyama-Nakashita A, Nakabayashi K, Li W, Ogawa M, Yamauchi Y, Preston J, Aoki K, et al (2008) The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. Plant J **55:** 526–542

Henderson IR, Jacobsen SE (2007) Epigenetic inheritance in plants. Nature **447:** 418–424

Hong F, Breitling R (2008) A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. Bioinformatics **24:** 374–382

Ideker T, Galitski T, Hood L (2001) A new approach to decoding life: systems biology. Annu Rev Genomics Hum Genet **2:** 343–372

Ishida T, Kurata T, Okada K, Wada T (2008) A genetic regulatory network in the development of trichomes and root hairs. Annu Rev Plant Biol **59:** 365–386

Iyer-Pascuzzi A, Simpson J, Herrera-Estrella L, Benfey PN (2009) Functional genomics of root growth and development in Arabidopsis. Curr Opin Plant Biol **12:** 165–171

Jiao Y, Tausta SL, Gandotra N, Sun N, Liu T, Clay NK, Ceserani T, Chen M, Ma L, Holford M, et al (2009) A transcriptome atlas of rice cell types uncovers cellular, functional and developmental hierarchies. Nat Genet **41:** 258–263

Kasschau KD, Fahlgren N, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Carrington JC (2007) Genome-wide profiling and analysis of Arabidopsis siRNAs. PLoS Biol **5:** e57

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. Plant J **50:** 347–363

Kwak SH, Schiefelbein J (2008) A feedback mechanism controlling SCRAMBLED receptor accumulation and cell-type pattern in Arabidopsis. Curr Biol **18:** 1949–1954

Levesque MP, Vernoux T, Busch W, Cui H, Wang JY, Blilou I, Hassan H, Nakajima K, Matsumoto N, Lohmann JU, et al (2006) Whole-genome analysis of the SHORT-ROOT developmental pathway in Arabidopsis. PLoS Biol **4:** e249

Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar

AH, Ecker JR (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell **133:** 523–536
Long TA, Brady SM, Benfey PN (2008) Systems approaches to identifying gene regulatory networks in plants. Annu Rev Cell Dev Biol **24:** 81–103
Morohashi K, Grotewold E (2009) A systems approach reveals regulatory circuitry for Arabidopsis trichome initiation by the GL3 and GL1 selectors. PLoS Genet **5:** e1000396
Ohtsu K, Smith MB, Emrich SJ, Borsuk LA, Zhou R, Chen T, Zhang X, Timmermans MC, Beck J, Buckner B, et al (2007) Global gene expression analysis of the shoot apical meristem of maize (Zea mays L.). Plant J **52:** 391–404
Popescu SC, Popescu GV, Bachan S, Zhang Z, Gerstein M, Snyder M, Dinesh-Kumar SP (2009) MAPK target networks in Arabidopsis thaliana revealed using functional protein microarrays. Genes Dev **23:** 80–92
Price ND, Shmulevich I (2007) Biochemical and statistical network models for systems biology. Curr Opin Biotechnol **18:** 365–370
Priest HD, Filichkin SA, Mockler TC (2009) cis-Regulatory elements in plant cell signaling. Curr Opin Plant Biol **12:** 643–649
Pruneda-Paz JL, Breton G, Para A, Kay SA (2009) A functional genomics approach reveals CHE as a component of the Arabidopsis circadian clock. Science **323:** 1481–1485
Qin Y, Leydon AR, Manziello A, Pandey R, Mount D, Denic S, Vasic B, Johnson MA, Palanivelu R (2009) Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. PLoS Genet **5:** e1000621
Raikhel NV, Coruzzi GM (2003) Achieving the in silico plant: systems biology and the future of plant biological research. Plant Physiol **132:** 404–409
Savage NS, Walker T, Wieckowski Y, Schiefelbein J, Dolan L, Monk NA (2008) A mutual support mechanism through intercellular movement of CAPRICE and GLABRA3 can pattern the Arabidopsis root epidermis. PLoS Biol **6:** e235
Schauer N, Fernie AR (2006) Plant metabolomics: towards biological function and mechanism. Trends Plant Sci **11:** 508–516
Schiefelbein J, Kwak SH, Wieckowski Y, Barron C, Bruex A (2009) The gene regulatory network for root epidermal cell-type pattern formation in Arabidopsis. J Exp Bot **60:** 1515–1521
Spencer MW, Casson SA, Lindsey K (2007) Transcriptional profiling of the Arabidopsis embryo. Plant Physiol **143:** 924–940
Tominaga R, Iwata M, Okada K, Wada T (2007) Functional analysis of the epidermal-specific MYB genes CAPRICE and WEREWOLF in *Arabidopsis*. Plant Cell **19:** 2264–2277
Vermeirssen V, Barrasa MI, Hidalgo CA, Babon JA, Sequerra R, Doucette-Stamm L, Barabasi AL, Walhout AJ (2007) Transcription factor modularity in a gene-centered C. elegans core neuronal protein-DNA interaction network. Genome Res **17:** 1061–1071
Wang S, Kwak SH, Zeng Q, Ellis BE, Chen XY, Schiefelbein J, Chen JG (2007) TRICHOMELESS1 regulates trichome patterning by suppressing GLABRA1 in Arabidopsis. Development **134:** 3873–3882
Watson AD (2006) Thematic review series: systems biology approaches to metabolic and cardiovascular disorders. Lipidomics: a global approach to lipid analysis in biological systems. J Lipid Res **47:** 2101–2111
Wu Y, Llewellyn DJ, White R, Ruggiero K, Al-Ghazi Y, Dennis ES (2007) Laser capture microdissection and cDNA microarrays used to generate gene expression profiles of the rapidly expanding fibre initial cells on the surface of cotton ovules. Planta **226:** 1475–1490
Yadav RK, Girke T, Pasala S, Xie M, Reddy GV (2009) Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. Proc Natl Acad Sci USA **106:** 4941–4946
Zhang X, Madi S, Borsuk L, Nettleton D, Elshire RJ, Buckner B, Janick-Buckner D, Beck J, Timmermans M, Schnable PS, et al (2007) Laser microdissection of narrow sheath mutant maize uncovers novel gene expression in the shoot apical meristem. PLoS Genet **3:** e101
Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, et al (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. Cell **126:** 1189–1201
Zhao M, Morohashi K, Hatlestad G, Grotewold E, Lloyd A (2008a) The TTG1-bHLH-MYB complex controls trichome cell fate and patterning through direct targeting of regulatory loci. Development **135:** 1991–1999
Zhao Z, Zhang W, Stanley BA, Assmann SM (2008b) Functional proteomics of *Arabidopsis thaliana* guard cells uncovers new stomatal signaling pathways. Plant Cell **20:** 3210–3226
Zhu JK (2008) Epigenome sequencing comes of age. Cell **133:** 395–397
Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S (2007) Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet **39:** 61–69