Annual Review of Plant Biology

Time-Based Systems Biology Approaches to Capture and Model Dynamic Gene Regulatory Networks

Jose M. Alvarez,[1,2] Matthew D. Brooks,[3] Joseph Swift,[4] and Gloria M. Coruzzi[5]

[1]Centro de Genómica y Bioinformática, Facultad de Ciencias, Universidad Mayor, Santiago, Chile

[2]Millennium Institute for Integrative Biology (iBio), Santiago, Chile

[3]Global Change and Photosynthesis Research Unit, US Department of Agriculture Agricultural Research Service, Urbana, Illinois 61801, USA

[4]Salk Institute for Biological Studies, La Jolla, California 92037, USA

[5]Center for Genomics and Systems Biology, Department of Biology, New York University, New York, NY 10003, USA; email: gc2@nyu.edu

Keywords

transcription factor, time-based genome-wide studies, systems biology, gene regulatory networks, transient regulatory events, dynamic network modeling

Abstract

All aspects of transcription and its regulation involve dynamic events. However, capturing these dynamic events in gene regulatory networks (GRNs) offers both a promise and a challenge. The promise is that capturing and modeling the dynamic changes in GRNs will allow us to understand how organisms adapt to a changing environment. The ability to mount a rapid transcriptional response to environmental changes is especially important in nonmotile organisms such as plants. The challenge is to capture these dynamic, genome-wide events and model them in GRNs. In this review, we cover recent progress in capturing dynamic interactions of transcription factors with their targets—at both the local and genome-wide levels—and using them to learn how GRNs operate as a function of time. We also discuss recent advances that employ time-based machine learning approaches to forecast gene expression at future time points, a key goal of systems biology.

Contents

1. INTRODUCTION

Unraveling the dynamic nature of gene regulatory networks (GRNs) is both a main promise and a key challenge of systems biology. The promise is the ability to exploit time-series data to learn the GRNs that drive dynamic genome-wide responses in a biological system. Doing so will allow us to "achieve the final aim of kinetic research; namely, to obtain knowledge of the nature of the reaction from a study of its progress," as stated in 1913 by Leonor Michaelis and Maud Menten (58, p. 8268; 85). In this review, we report the promise, progress, and challenges of capturing and modeling dynamic events in GRNs.

GRN: gene regulatory network

TF: transcription factor

Arguably, the most significant challenge is that our current understanding of how transcription factors (TFs) direct GRNs comes from biochemical investigations that favor stable TF–target gene interactions but largely miss transient ones. This review highlights studies which exploit time-series data that capture dynamic TF–target gene binding and/or regulation events, used to build or validate predictive GRN models. These time-based studies have revealed that transient interactions in GRNs—important for dynamic responses to their environment—have largely been missed by traditional methods that favor stable TF–target interactions.

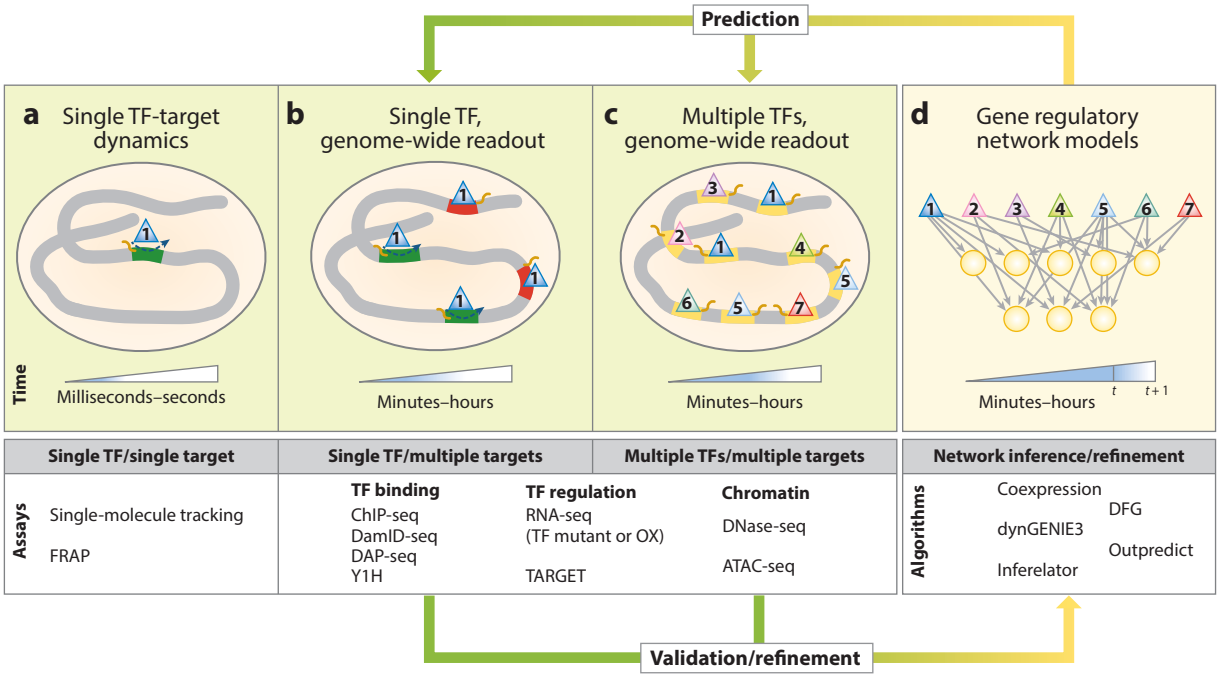

Figure 1

Overview of approaches to capture TF-target interaction dynamics. (*a*) Single TF–target gene dynamics. For a single TF (*numbered triangle*), single-molecule tracking technology captures TF–target gene binding dynamics within a time frame of milliseconds to seconds. See **Figure 2** for more detail. (*b*) Single-TF genome-wide target readout. ChIP-seq, DamID-seq, and DAP-seq assays capture the genome-wide relevance of single-TF binding within minutes to hours. RNA-seq and TARGET can capture TF-regulated genes genome-wide for a single TF. See **Figure 3** for more detail. (*c*) Multiple-TF genome-wide target readout. The genome-wide relevance of multiple TFs can be captured in the same experimental setup by chromatin profiling assays such as DNAse-seq and ATAC-seq. Multiple TFs can also be studied in parallel by ChIP-seq or TARGET. See **Figure 4** for more detail. (*d*) Time-based gene regulatory network modeling and validation. Machine learning algorithms can identify TF–target causality by time-series data. See **Figure 5** for more detail. Such time-based methods can also be used to infer gene expression at future time points ($t + 1$). Experimental data from genome-wide studies of TF–target binding or regulation can be used as priors or to refine the network models using precision/recall analysis. Abbreviations: ATAC-seq, assay for transposase-accessible chromatin using sequencing; ChIP-seq, chromatin immunoprecipitation sequencing; DamID-seq, DNA adenine methyltransferase identification sequencing; DAP-seq, DNA affinity purification and sequencing; DFG, dynamic factor graph; FRAP, fluorescence recovery after photobleaching; TARGET, transient assay reporting on genome-wide effects of TFs; TF, transcription factor.

This review focuses on recent advances in time-based studies that capture and model the dynamics of TF–target interactions in plant GRNs. First, we review new technologies for capturing the dynamics of TF–target interactions in vivo—at single target genes, drawing mostly from studies in animals (**Figure 1**). Second, we review novel cell-based TF-perturbation methods in *Arabidopsis* that enable the capture of transient TF–target interactions genome-wide, and evaluate their impact on mediating rapid transcriptional responses (6, 11, 96). The increasing availability of new experimental technologies for capturing TF–target interactions allows multiple assays to be performed on a set of TFs, rather than individual TFs. These new data sets are facilitating the construction of validated hierarchical GRNs on the basis of TF binding (92, 113) or TF-mediated gene expression (20) (**Figure 1**). Indeed, such analysis can identify influential TFs at the top of a temporal TF cascade whose modifications can have dramatic effects on downstream gene regulation and the resulting phenotypes (20, 113, 128).

SMT: single-molecule tracking

Importantly, the decreasing cost of sequencing technologies has allowed researchers to generate high-resolution time-course data sets for TF–target gene interactions genome-wide (20, 113, 128). These data sets are aiding the development of new computational time-based methods that can construct predictive GRNs. For example, machine learning strategies have been used to build temporal biological networks (20, 128). Crucially, including time-series data as a dimension to interrogate GRNs can even allow one to infer gene expression states at future time points, a main goal of systems biology (25, 128) (**Figure 1**).

2. SINGLE TRANSCRIPTION FACTOR–TARGET GENE DYNAMICS

2.1. Molecular Mechanisms of Transcription Factor–Target Dynamics and Its Effect on Gene Regulation

TF-binding sites at bona fide target genes are sparse in the genome compared with the number of nonspecific TF-binding sites (55). How, then, do TFs find and bind to their specific sites in the genome to coordinate rapid transcriptional responses? And how does this process play out dynamically in real time? Rapid changes in gene expression of multiple target genes simultaneously are governed by the kinetic properties of TFs, which include their ability to rapidly diffuse within the nucleus and their propensity to transiently bind to their genome-wide targets.

Methods that track the movement of individual TF molecules in vivo have provided insights into how TF kinetics affect target binding and gene transcription. To date, methods for capturing the dynamics of TF–target interactions in situ have been conducted mostly in animals (e.g., *Drosophila*), and they have only recently been applied in plants (26). Such approaches, including single-molecule tracking (SMT), have revealed that TFs do not bind permanently to their target *cis*-regulatory sites (70, 75, 89, 94). These in vivo assays, which typically track TF binding at a single reporter gene, show that the residence time of a TF at a specific gene target lasts from hundreds of microseconds to several seconds—just long enough to initiate transcription (70, 75, 89, 94) (**Figure 2*a***). For example, the residence time for the mammalian TF p53 at its *cis*-regulatory sites on a gene target is approximately 3.5 s, while for the glucocorticoid receptor (GR) it is 8.1 s (89). TFs can also bind nonspecifically to DNA targets at noncanonical sites; however, this association is weak, resulting in residence times that are an order of magnitude shorter (75). Notably, other techniques indicate that binding times could be much longer. For example, TF–DNA binding competition assays report that TF binding to a target gene can last for minutes or even hours (46, 73).

Given that TFs bind to a target gene transiently in vivo, with residence times likely ranging from milliseconds to minutes, only a small proportion of TFs within a nucleus are in the bound state at any given time (8). Indeed, estimates based on SMT of steroid receptors in mice place the fraction of TFs bound to DNA between 5% and 10% (94). Similarly, at any given moment, not all available *cis*-regulatory sites for a particular TF are bound, especially when the number of *cis*-regulatory sites vastly outnumbers the number of TF molecules (19). The effect this has on transcriptional output is that the number of transcript molecules arising from a target locus is not constant over time (16, 17, 70). Instead, it is stochastic, with transcription occurring in bursts whose frequencies are dependent on the number of TF–target interactions as well as on the availability of TF-related cofactors (12, 61, 66, 115). Stochasticity in transcriptional output not only has been visualized at the single-molecule level (70) but also is reflected in the heterogeneity of genome-wide expression profiles within single cells, as found by single-cell RNA sequencing (RNA-seq) in plants (30) and other eukaryotes (22, 57).

Gene expression levels often change over time in response to environmental or developmental cues. At the molecular level, a change in the number of transcripts produced by a locus can

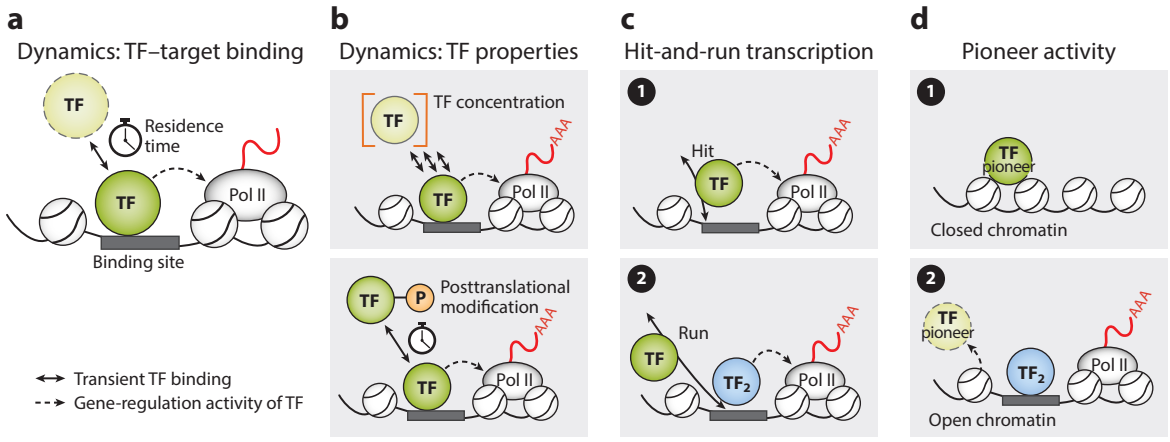

Figure 2

The molecular mechanisms underlying TF–target binding dynamics. (*a*) TFs interact dynamically with their binding site in a target gene; binding times last from milliseconds to minutes. Both the number of TF–target interactions and the time the TF spends residing at the binding site can affect transcription output. (*b*) An increase in the concentration of a TF can increase the number of TF–target binding events. Similarly, posttranslational modifications to a TF can increase residence time. Both scenarios result in an increase in transcriptional output. (*c*) The hit-and-run transcription model. Here, a TF binds transiently to initiate transcription (the hit, ❶), while its dissociation (the run, ❷) allows transcription to be maintained by TF_2s (see **Figure 3**). (*d*) Some TFs display pioneer activity, where their binding to closed chromatin (❸) alters its accessibility to allow TF_2 binding (❹). Abbreviations: Pol, polymerase; TF, transcription factor, TF_2, secondary TF.

be achieved through either a change in the TF's residence time at a gene target or a change in the number of TF–target interactions (**Figure 2*b***). TF–target residence times can be affected by posttranslational modifications (75) or the presence of TF-binding partners (51, 75, 94), and an increase in the TF residence time at a gene target can lead to an increase in transcriptional output. Likewise, a change in the concentration of a TF can lead to a proportional change in the number of TF–target binding events that initiate transcription of a target gene (8, 19) (**Figure 2*b***). While shorter TF residence times on a target gene are associated with weaker transcriptional output [an effect referred to as treadmilling (73)], such rapid on–off TF–target gene binding is thought to help reduce noise within gene expression systems (8, 45).

Notably, while elegant methods to track TF–target binding in vivo have provided detailed insight into the dynamics of TF–target binding, such assays typically rely on a single reporter gene. Consequently, techniques such as SMT cannot capture the impact that dynamic TF–target interactions have genome-wide. Additionally, they do not show how multiple TFs combine to regulate gene expression, which is required to model causal GRNs. These topics are discussed at length below.

Other molecular mechanisms exist that allow an increase in transcriptional output to occur without an increase in TF concentration. These mechanisms fall into two broad categories. The first category aims to keep TFs localized as closely as possible to their binding site at a target gene. This category includes (*a*) formation of large multimolecular assemblies that localize TFs to particular *cis*-regulatory locations (38, 52, 88), (*b*) tethering of TFs by the target gene's RNA product (112), and (*c*) blocking of TF nuclear export (83). The second category aims to increase the number of transcripts per TF–target binding event. For example, in so-called hit-and-run transcription (discussed at length in the next section) (84), a TF need only bind its *cis*-regulatory element briefly (the hit) to initiate transcription that continues after the run, where secondary TFs (TF_2s)are likely responsible for maintaining transcriptional output (6, 24, 96, 127) (**Figure 2*c***).

Treadmilling: short TF residence time that is associated with weak transcriptional output

Hit and run: a TF hit can initiate a transcriptional complex, enabling transcription to continue after the initiating TF is no longer bound

2.2. The Role of Chromatin in Transcription Factor–Target Gene Dynamics

Pioneer TF: a TF that binds and shifts heterochromatin to more accessible chromatin, allowing nonpioneer TFs to bind

TAD: topologically associated domains

DAP-seq: DNA affinity purification sequencing

The genome is not a flat, naked DNA template poised for TF–target binding but rather a three-dimensional landscape that TFs must navigate. The three-dimensional organization of chromatin affects TF–target gene binding dynamics. A key example is the way condensed heterochromatin limits the ability of TFs to access DNA. Typically, nongenic regions of chromosomes (such as centromeres), as well as genic regions that do not play a role in cell function, are packaged tightly into histones. Condensing large portions of the genome minimizes the nonspecific TF–DNA binding events that can occur and promotes TF binding to the *cis*-regulatory sites that remain accessible. Importantly, certain so-called pioneer TFs, such as Zelda in *Drosophila* (38), play a part in determining the developmental timing of which regions of the genome are open and which are closed. Such pioneers interact with heterochromatin, shifting histones to expose genic regions, in turn allowing nonpioneer TFs to bind and regulate gene expression (68) (**Figure 2*d***). Pioneer TF activity appears particularly important during early organismal development and has been characterized largely within the developmental programs of mammals and *Drosophila* (68). However, pioneer TFs have also been found to be active during early seed development in plants (120), indicating that this class of TFs also exists in the plant kingdom (102).

Areas of open chromatin can also hold additional structural features that can impact TF–target binding. Topologically associated domains (TADs) are segments of the genome, typically in the range of tens to hundreds of kilobases, that fold upon themselves to create localized genomic compartments (119). Compartmentalizing portions of the genome into TADs affects gene regulation. For example, within TADs, genes often display similar levels of gene expression (99), and the effect of enhancers appears not to extend beyond TAD borders (119). At present, there is weak evidence that TADs exist in *Arabidopsis*; not only does its genome lack a close homolog of the CTCF gene required for TAD formation (35, 119), but also chromosomal chromatin capture experiments fail to detect such domains. However, growing evidence indicates that TADs exist in plants with larger genomes, such as wheat (27), cotton (129), and maize (80).

3. TRANSCRIPTION FACTOR–TARGET GENE DYNAMICS AND GENOME-WIDE EXPRESSION

As described above, TF–target gene dynamics are an intrinsic property of gene regulation. The fine-scale time resolution of TF–target gene binding has revised previous notions of stable TF–target interactions of the pioneer TF Zelda and the TF Bicoid—which are now known to interact highly transiently with TF-dependent transcription sites in developing *Drosophila* embryos (88). However, while SMT studies can detect transcriptional dynamics with a single target gene within seconds or milliseconds; they cannot capture the genome-wide TF–target interactions.

Capturing dynamic interactions on a genome-wide scale has been a technical challenge, largely because prevailing techniques to capture TF-bound targets genome-wide in vivo or in vitro, such as chromatin immunoprecipitation (ChIP) and DNA affinity purification sequencing (DAP-seq) (92), respectively, reflect only a snapshot of the most stable TF interactions at a single time point (28, 67, 125). Time-series ChIP-seq experiments, such as those conducted on the TF EIN3 in *Arabidopsis*, have contributed to our knowledge of TF–target dynamics in planta. For many EIN3-bound and ethylene-responsive genes, EIN3 binding peaked at 4 h after the ethylene treatment and weakened after 12 h. In contrast, EIN3-induced target genes showed sustained levels of transcripts during the full course of the 24-h ethylene treatment, despite the weakened binding of EIN3 after 4 h (23). Although time-series ChIP is a suitable approach to the evaluation of TF–target gene binding within hours, it presents drawbacks when it comes to analyzing rapid changes

that could occur within seconds to minutes. This is in part because the time required to fix TF–DNA complexes in whole tissues for ChIP requires a minimum of ~15 min, which is usually longer than gene expression changes in response to acute signals. In addition, ChIP assays—which are based on biochemical associations—favor stable TF–target binding and miss transient TF–target interactions (72).

TARGET: transient assay reporting on genome-wide effects of TFs

DamID-seq: DNA adenine methyltransferase identification coupled to sequencing

Studies exploiting the fact that plant cell–based TF-perturbation assays have helped to capture early and transient TF–target binding events that are missed in planta (6, 96). One such example is the temporal TF perturbation assay called TARGET (transient assay reporting on genome-wide effects of TFs). In TARGET, a chosen TF is transiently expressed in isolated plant cells (protoplasts) as a TF–GR fusion protein (TF::GR) (11, 20, 96). To temporally control TF action, transfected root cells transiently expressing TF::GR are sequentially treated with cycloheximide (±CHX) to block regulation of TF_2 targets as well as dexamethasone (±DEX) to induce TF nuclear entry (11). Genes regulated by DEX-induced TF::GR nuclear import are deemed direct targets of the TF, since +CHX pretreatment blocks translation of downstream regulators (11). Several studies using plant cell–based TARGET have shown that direct TF-regulated genes are early and transiently bound targets of certain TFs (6, 34, 96). Although TARGET is a suitable means of studying the importance of binding time of direct transcriptional regulation, a caveat is that it may not necessarily provide an authentic biological context because it is performed in isolated plant cells (e.g., protoplasts) (6, 34, 96) (**Table 1**).

3.1. Hit-and-Run Model of Transcription and Its Rapid Effect on Transcriptome Reprogramming

The genome-wide relevance of transient TF–target interactions was initially demonstrated for bZIP1 by use of the plant cell–based TF TARGET (34, 96). The early and transiently bound genome-wide targets of bZIP1 were captured in TARGET by combining TF-mediated changes in gene regulation with time-series TF-binding data using micro-ChIP-seq (95, 96). These studies provided the first genome-wide evidence for a hit-and-run transcription model (24, 127), initially hypothesized in 1988 (103) but validated only on a single-gene basis in animals (39, 84) (**Figure 2*c***). The hit-and-run model posits that a TF trigger (the hit) can organize a stable transcriptional complex, including the recruitment of other TFs, so that transcription can continue even after the initiating TF is no longer bound (the run) (24, 34, 96, 127) (**Figure 2*c***). Time-series ChIP-seq experiments conducted 1–3 min following controlled TF nuclear import revealed that bZIP1 bound transiently specifically to promoters of early nitrogen-response genes (96). Moreover, 4-thiol-uracil labeling of de novo TF-initiated transcripts was used to affinity-capture bZIP1-initiated mRNAs, demonstrating that transiently bound bZIP1 targets were actively transcribed at times when the TF was no longer bound (34).

Transient binding of a TF to its regulated genome-wide targets was also recently captured for NIN-LIKE PROTEIN7 (NLP7) (6), an early-acting regulator of nitrogen signaling in plants (74, 83). One study demonstrated early and transient binding of NLP7 to its regulated targets by integrating fine-scale time-series ChIP with NLP7-dependent genome-wide regulation data (6), using the plant cell–based TARGET system. This study showed that highly transient NLP7 genome-wide targets were TF regulated in response to NLP7 nuclear import but that binding of NLP7 to these targets was not detected by ChIP at any time point tested (5–180 min) (6). To capture the highly transient NLP7 target genes that defied biochemical assays such as ChIP, the authors adapted the DNA adenine methyltransferase identification (DamID) method such that a TF fusion to a DNA adenine methyltransferase is coupled with sequencing (DamID-seq) (6, 7). In

Table 1 Overview of methods used to capture, model, and validate dynamic gene regulatory networks

a Single TF–target gene dynamics			
Transcription model	**Technique(s) used**	**Organism**	**Reference(s)**
Treadmilling	Competition ChIP	*Saccharomyces cerevisiae*	73
Hit-and-run	FRAP	*Mus musculus*	84
TF pioneers	Single-molecule tracking	*Drosophila melanogaster*, *Arabidopsis thaliana*	68, 120
Transient TF binding	Single-molecule tracking	Human, *Drosophila melanogaster*	38, 88, 89
Advantages: high temporal resolution, live-cell measurements		Drawbacks: miss genome-wide TF targets needed for GRN models	
b TF–target dynamics genome-wide			
Transcription model	**Technique(s) used**	**Organism/pathway**	**Reference(s)**
Hit-and-run (rapid responses)	TARGET and time-series micro-ChIP	*Arabidopsis thaliana*/nitrogen	34, 96
	TARGET, time-series micro-ChIP and DamID	*Arabidopsis thaliana*/nitrogen	6
Hit-and-run (chronic responses)	ChIP	*Arabidopsis thaliana*/heat shock	69
TF dynamics in hormone responses	Time-series ChIP	*Arabidopsis thaliana*/ethylene	23
TF binding versus regulation	TARGET and ChIP	*Arabidopsis thaliana*/nitrogen, ABA	21, 113
Advantages: can identify stable and transient TF–target interactions genome-wide		Drawbacks: TARGET uses protoplasting step and cell populations	
c The hierarchy of temporal GRNs			
Time-based GRN feature	**Characteristics**	**Result**	**Reference(s)**
Feed-forward loop	Coherent: TF_1 and TF_2 have same effect on target gene	Sign-sensitive delay in target induction or in turning off response	100
	Incoherent: TF_1 and TF_2 have opposite effects on target gene	Rapid and transient target induction (pulse-like behavior)	77
Feedback loop	TF_1 regulated by downstream TF_2 targets of TF_1	Generates oscillations in target expression	44, 108
Temporal TF hierarchy	Interactions between ABA-responsive TFs and targets revealed by time-series transcriptomics and ChIP	Interactions between 21 TFs and 3,061 ABA-responsive genes; validated two new regulators of ABA (DIG1 and DIG2)	113
	Interactions between nitrogen-responsive TFs and targets revealed by time-series transcriptomics and TARGET	Defined temporal hierarchy of 172 TFs and 2,174 targets in shoots and 145 TFs and 1,458 targets in roots; validated a new TF hub (CRF4) in nitrogen signaling	20, 128
Advantages: breaking networks into well-understood components		Drawbacks: difficult to validate the temporal role of such TF–TF network motifs in planta	

(*Continued*)

Table 1 (*Continued*)

d Modeling and validating time-based GRNs			
GRN modeling approach	**Characteristics**	**Examples**	**References**
Lagged correlation	Accounts for lag-lead effects in time-series data by including an offset in one variable	Temporal clustering by affinity propagation, light responses in Cyanobacteria	60, 104
Linear regression models	Used in ML methods that assume a linear relationship between TFs and target genes	Dynamic factor graph, Inferelator	20, 65, 128, 130
Nonlinear regression models	Used in ML methods, such as random forests, to account for more complex nonlinear relationships between TFs and target genes	GENIE3 (dynGENIE3), OutPredict	25, 82, 130
Network inference	Prediction of regulatory interactions between TFs and target genes	Dynamic factor graph, Inferelator, GENIE3, OutPredict	20, 25, 65, 82, 128, 130
Forecasting	Prediction of target gene expression values at future time-points	OutPredict, dynGENIE3	25, 41
Advantages: Time-series data can reveal casual interactions that cannot be identified from steady-state data. Regression models allow forecasting of gene expression in untested conditions, in addition to GRN reconstruction.		Drawbacks: Selection of the best method often involves testing many algorithms on a data set. Validation data for network inference in plants are relatively sparse, especially for TF-regulated edges.	

Abbreviations: ABA, abscisic acid; ChIP, chromatin immunoprecipitation; dynGENIE3, dynamical GENIE3; FRAP, fluorescence recovery after photobleaching; GRN, gene regulatory network; ML, machine learning; TARGET; transient assay reporting on genome-wide effects of TFs; TF, transcription factor.

DamID-seq, the Dam::TF fusion protein leaves a methylation mark on promoters touched, even transiently, by the TF (7). By adapting the DamID-seq approach to the plant cell–based TARGET system, the authors captured NLP7 binding to highly transient NLP7 targets that defied detection even by time-series ChIP (6).

Notably, DamID detects any binding events that occurred during the time frame of the experiment; thus, it is not limited by the snapshot problem of ChIP-seq and other antibody-based techniques (7, 107, 110). In this context, time-series ChIP-seq and DamID-seq are sensitive approaches to the capture of NLP7-regulated genes either in root cells or in planta (6). However, the number of NLP7-bound genes identified by either ChIP or DamID greatly exceeds the number of genes directly regulated by NLP7. Indeed, several studies have demonstrated TF binding at a large number of sites, many of which cannot be clearly connected with target gene regulation (76). The apparently low specificity for ChIP (i.e., the large proportion of TF-bound targets that are not TF regulated) holds for DamID (6, 110), supporting the notion that TF binding is necessary but not sufficient for gene regulation. One proposed explanation for widespread TF binding is the presence of nonfunctional binding sites that serve no biological purpose (76). Alternatively, TFs can bind to many low-affinity sites in the genome and contribute to gene expression at levels that are low but sufficient to allow evolutionary conservation (121). Thus, the actual specificity is difficult to estimate, since TF-bound genes that are not transcriptionally activated may be poised and waiting for cofactors or other signals to induce gene expression (23, 76).

NLP7-dependent de novo transcribed targets identified using 4-thiol-uracil affinity labeling showed that highly transient NLP7 targets are transcriptionally active even after NLP7 is no

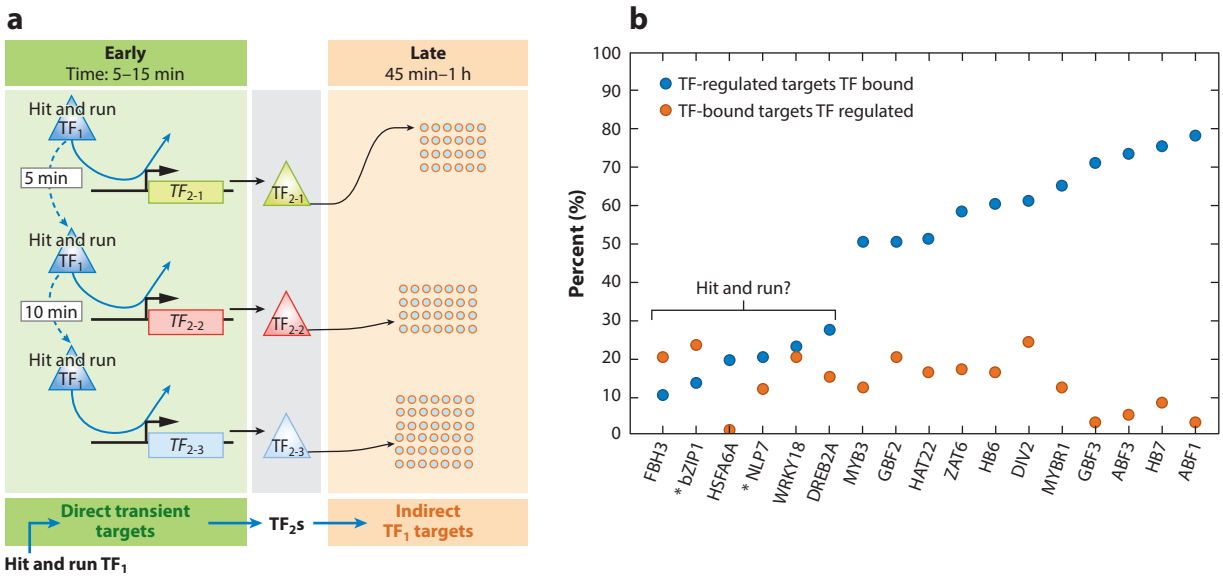

Figure 3

Transient binding of hit-and-run TFs highlights the discord between TF-binding and TF-regulation data. (*a*) Transient TF–target interactions by hit-and-run TF_1s (e.g., bZIP1 and NLP7) initiate early transcriptional responses to an environmental cue (e.g., nitrogen signals) (6, 96). The transient interactions of a hit-and-run TF_1 enable it to rapidly activate TF_2s, leading to an increase in the transcriptional output in a short period of time. TF_2s amplify the transcriptional cascade by regulating indirect targets of the TF_1. (*b*) The overlap of TF-binding and TF-regulation data for 17 TFs (21) in the ABA (113) and nitrogen signaling (6, 96) pathways. Overall, TF regulation is a good indicator of TF binding (*blue dots*). TF binding is a poor indicator of TF regulation (*orange dots*). Six TFs show a low percentage of TF-regulated targets that are TF bound (*first six blue dots*), which may act as hit-and-run TFs. Two of these were confirmed as hit-and-run TFs using time-series data (*bZIP1 and *NLP7) (6, 96). Abbreviations: ABA, abscisic acid; TF, transcription factor; TF_1, primary TF; TF_2, secondary TF.

longer bound (6). This result confirmed that transient interactions of NLP7—missed even by time-series ChIP—initiate active transcription, according to the hit-and-run transcription model. Importantly, this study showed that these transient NLP7 target interactions activate early N-responsive secondary TFs (TF_2s), which directly regulate the expression of late N-responsive genes. Moreover, the direct TF_2 targets of NLP7 themselves directly regulate a large number of genes indirectly controlled by NLP7 *in planta* (6) (**Figure 3*a***).

What is the purpose of hit-and-run transcription? This model might ensure that a small number of TF molecules can rapidly affect a large number of target genes in a very short period (5–90 min) (24, 127), which would allow an organism to rapidly adapt to a change in its external or internal nutritional status. This notion is supported by the finding that the transient hit-and-run targets of both bZIP1 (34, 96) and NLP7 (6) are enriched in genes responding very early to nitrogen signals (e.g., in 3–5 min) (4, 6, 34, 65, 96, 128). In this context, rapid and transient interactions of a hit-and-run TF can directly activate additional downstream regulatory circuits, amplifying transcriptional output via TF_2s, as shown for NLP7 (6) (**Figure 3*a***). The data supporting the hit-and-run model distinguish it from other dynamic transcription models such as treadmilling (**Table 1**), where transitive binding to targets results in a lower level of activation (73).

The hit-and-run model of transcription described for chronic responses, such as heat shock stress (69), may differ from rapid responses to nutrient signals (6, 96) (**Table 1**). For chronic

responses, lasting changes in chromatin modifications allow plants to form stress memory (32, 69). The transcriptional memory following heat shock stress correlates with histone H3 lysine 4 di- and trimethylation at memory-related loci (69). Here, the hit-and-run TF HSFA2 transiently associates with gene loci, leading to sustained transcriptional activity and accumulation of histone H3 lysine 4 methylation after TF binding has decreased (69). This mechanism allows plants to maintain the activity of heat shock memory–related genes. However, the time frame of the heat shock response is 4 to 52 h (69), which differs from rapid nitrogen-nutrient responses within 5 to 90 min (65, 128). Whether the maintenance of transcriptional activity in response to the rapid nitrogen nutrient signals is associated with histones and chromatin modifications remains unknown. Considering that nitrogen treatments do not produce changes in chromatin accessibility in a comparable time frame (4), it is likely that the sustained transcription of targets that were hit by the TF is achieved by TF-interacting partners, rather than histone or chromatin modification. This hypothesis is supported by enrichment of *cis* elements for partner TFs in the promoters of transiently bound TF targets of both bZIP1 and NLP7 (6, 96). That said, models for hit-and-run transcription invoke pioneer activity as well as changes in chromatin state and partner TFs (24, 115, 127). Notably, hit-and-run transcription has also been invoked for models of gene repression (107).

The extent of transient TF–target interactions in hit-and-run transcription also suggests a TF activity model similar to that of a catalyst (127), where a small number of TF molecules can rapidly affect a large number of target genes by acting catalytically. The concept of TFs acting as a catalyst was recently supported in experiments on TGA1, a master TF that acts as a regulator of nitrogen dose response (117). By modeling genome-wide transcriptional responses to nitrogen dose as a function of time, the experimenters used the Michaelis–Menten model (1, 117), developed to explain enzyme kinetics (58), to describe the dynamics of plant transcriptomic responses as well as changes in nitrogen dose–related growth. Increased levels of TGA1 can enhance the maximum rate of gene expression, which agrees with the Michaelis–Menten model, wherein an increase in the amount of catalyst produces higher rates of the reaction. While TGA1 is a high-level regulator of nitrogen responses, targeting the expression of 92 other TFs, whether TGA1 acts as a hit-and-run TF remains to be determined (117).

3.2. Transcription Factor–Target Gene Binding Often Does Not Result in Transcription Factor Regulation

ChIP-seq has been widely used to capture TF binding to targets that control responses to a myriad of hormonal and environmental signals in plants (123). However, due to the biochemical limitations of ChIP (which favors stable TF–target gene interactions) and the limited number and frequency of the time points that can be analyzed, the prevalence of transient TF–target binding via a hit-and-run mechanism may be underappreciated. For example, a paradox in the biological community is that in genome-wide studies of TF–target binding (across plants, yeast, and other animals), the percent of TF-bound genes that are TF regulated is typically low (118). For example, for NLP7, the proportion of TF-bound genes (assayed by ChIP) that are TF regulated in planta is ~20% (83). The elusive 80% of genes that are regulated by NLP7 but not stably bound (e.g., by ChIP) likely involve transient TF–target interactions—a hypothesis supported by a study by Alvarez et al. (6).

To address this issue, Brooks et al. (21) explored the relationship of TF binding versus TF regulation for multiple TFs, including ones in the abscisic acid (ABA) signaling pathway (113). This analysis revealed three interesting findings:

ABA: abscisic acid

1. Stable TF binding is a poor indicator of TF regulation. For all 17 TFs examined, only 2–20% of TF-bound genes are TF regulated (**Figure 3*b***).
2. TF regulation is a good indicator of TF binding for the majority of TFs examined. For 11 out of 17 TFs, 50–80% of TF-regulated genes are stably TF bound, as detected by ChIP (**Figure 3*b***).
3. However, for 6 out of 17 TFs, only a small percentage of TF-regulated genes are TF bound (10–25%) (**Figure 3*b***). This set of 6 TFs includes 2 confirmed hit-and-run TFs (NLP7 and bZIP1) (6, 96) and an additional 4 candidate hit-and-run TFs (FBH3, HSFA6A, WRKY18, and DREB2A) (**Figure 3*b***).

Indeed, this study (21) and others suggest that TF–target binding by ChIP is often a poor prediction of TF regulation (43, 83, 118) and reflects only a snapshot of the most stable TF interactions at the time point analyzed, most of which are not associated with regulation (28, 67, 72, 125).

4. THE HIERARCHY OF TEMPORAL GENE REGULATORY NETWORKS

Because plants are sessile, they must be able to mount a rapid transcriptional response to a change in their environment. Creating such a rapid genome-wide transcriptional response is not the work of any single TF but rather results from the combined action of many TFs working in a temporal cascade. For this reason, understanding the behavior of TF networks over time can offer important insights into how environmental cues trigger differential gene expression. Indeed, studying how TFs work as an integrated system over time can reveal emergent properties of gene regulation that cannot be detected by studying a single TF.

4.1. Network Architecture of Transcription Factors Enables Dynamic Gene Regulatory Network Responses

In plants, a change in a key input—such as light quality, nutrient availability, or water—can result in a change in the expression level of thousands of genes (48, 109, 128). Such large changes are typical when complex physiological responses—such as a change in growth rate or transition to flowering—are required (126). To ensure the proper regulation of numerous genes, some TFs within signaling networks act as TF hubs. Specifically, the number of targets each TF has within a signaling response (known as out-degree) is thought to be scale free and follow a power-law distribution, with only a few TFs having many connections (2). Such TF hubs can regulate the expression of hundreds to thousands of genes (97), thus playing an outsized role in signal propagation, and their perturbation can disproportionately affect phenotype (128).

Within GRNs, a target gene's expression is often regulated by more than one TF. Indeed, the number of TFs that regulate a particular gene (known as in-degree), like out-degree, also appears to be scale free (2). A gene that has multiple TF regulators is likely to be more robustly expressed within the signaling network (113). Importantly, when a gene target is regulated by multiple TFs, many of those TFs can be homologous. For example, the TF homologs MYC2 and MYC3, involved in light signaling, share many of the same gene targets (131). The same is true for the TGA1 and TGA4 homologs in nitrogen signaling (117). Such redundancy creates resilient networks; not only does such redundancy provide an extra copy of a TF if one is mutated (5), but also multiple homologous TFs likely normalize the number of transcriptional bursting events, leading to better buffered-expression patterns (115).

4.2. Generating Temporal Transcriptional Networks from Time-Series Data

DHS: DNase I hypersensitive site

Footprint: TF-bound sequence protected from DNase I digestion

While time-series data will reveal when a given TF transcriptionally responds to a signal, it is also important to understand the regulatory relationships among TFs themselves. To this end, it is necessary to identify TF–target gene interactions for not only one TF but multiple TFs involved in a signaling response. For example, the combination of an ABA-response time course with ChIP-seq on 21 ABA-responsive TFs allowed investigators to use interactions between each of these 21 TFs and its ABA-responsive target genes to uncover a TF hierarchy for ABA signaling (113).

Another study delineated the temporal hierarchy of TF–target gene interactions governing the response of *Arabidopsis* to nitrogen signals by using a fine-scale time-series experiment that monitored the transcriptome-wide nitrogen response in both shoots and roots (128). Using a just-in-time analysis that binned genes on the basis of the first time point when their nitrogen response passed a threshold of 1.5-fold, the authors of this study identified a temporal TF cascade in each organ (128). Importantly, the TFs mediating this nitrogen-response cascade were validated with the cell-based TARGET, which can identify direct regulated targets by using controlled nuclear entry of the TF. This approach validated the role of 7 TFs in shoots (128) and 33 TFs in roots (20) as major controllers of the nitrogen response in each organ in planta.

Hierarchical TF networks can also be constructed in a single experimental setup through genome-wide mapping of putative TF-binding sites in regions of accessible chromatin. This can be done using the DNase I enzyme, which cleaves accessible chromatin regions known as DNase I hypersensitive sites (DHSs). These sites are typically associated with active transcriptional regulation, RNA polymerase II engagement, and regions where TFs bind (133). The resulting DNA fragments of DNase I cleavage are sequenced (DNase-seq) to map the DHSs (116, 133). Performing DNase-seq at a high sequencing depth enables single-base-resolution identification of TF-bound sequences, since TF occupancy blocks DNase I digestion, leaving a footprint (91). Combining footprinting with known TF motifs allows generation of transcriptional networks that contain all of the bound TFs thus detected. When done as a time course, DNase-seq can uncover TF binding events in response to a signal. To this end, Sullivan et al. (116) analyzed the regulatory landscape dynamics of *Arabidopsis* during heat shock and photomorphogenesis to reveal the topology of TF–TF regulatory interactions. They found that the light-responsive network is enriched in autoregulatory loops and that TF–TF interactions are highly dynamic. These authors showed that during the heat shock response many heat shock factor TFs respond to the treatment and form a densely connected network including several novel feedback loops (116).

4.3. Complex Hierarchical Networks Are Built from Simple Components

GRNs, like many other complex networks, consist of smaller, reoccurring gene regulatory interactions that occur more often than expected than if the network were random (47). These patterns, called network motifs, enable the organism both to carry out temporal gene expression responses to a stimulus and to filter stimuli so that a gene is transcribed only under desired conditions (3). Originally characterized in single-celled organisms such as yeast (87) and *Escherichia coli* (79), network motifs are fundamental components of plant GRNs (90).

Feed-forward loops, in which two TFs work together to dynamically regulate target gene expression, are an example of such network motifs (3). In an incoherent feed-forward loop, the first TF [$TF_{2\text{-}2}$ (**Figure 4**)] activates a gene target, as well as activating a second TF [$TF_{3\text{-}6}$ (**Figure 4**)]. This second TF, once expressed, represses the target gene (**Figure 4**). This network motif enables an early and rapid pulse of target gene expression. For example, an incoherent feed-forward loop controls expression of the important nitrate transporter NRT2.1 in *Arabidopsis*.

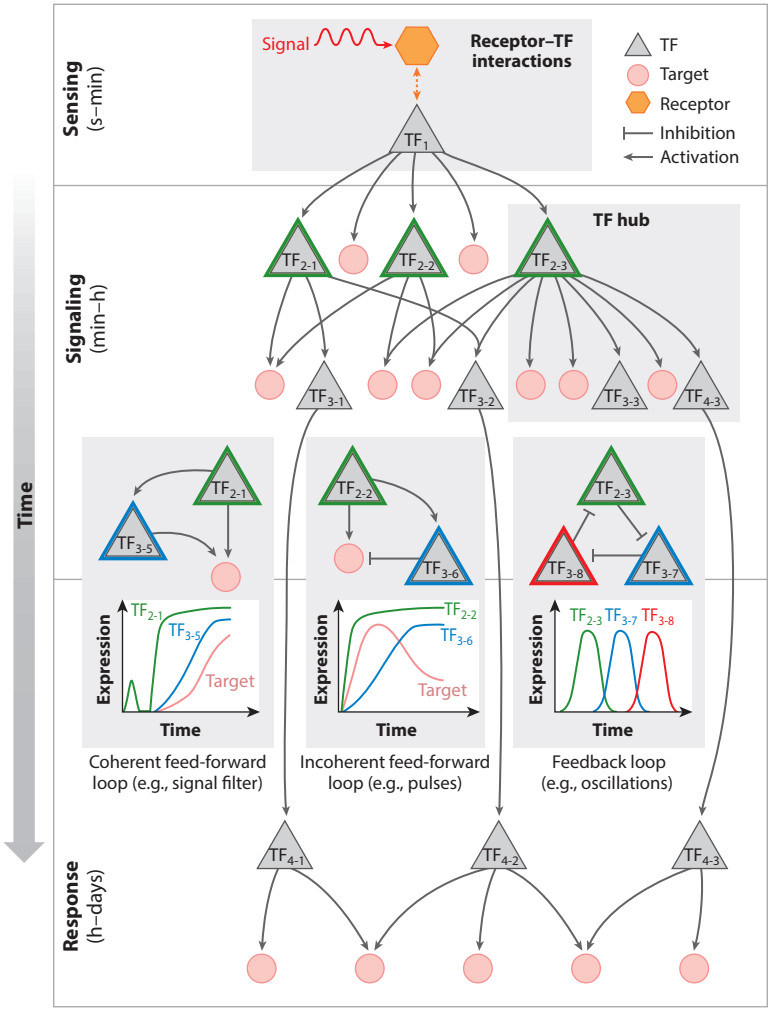

Figure 4

General features of hierarchical dynamic GRNs. GRNs, or transcriptional signaling cascades, are initiated through receptor–TF interactions, which can occur within seconds to minutes of an environmental or developmental signal. Within GRNs, TF hubs play an outsized role in regulating downstream target genes and secondary TFs. Gene regulatory interactions within GRNs can include feed-forward loops and feedback loops, which can allow target genes to either pulse or oscillate in their expression pattern (3). Abbreviations: GRNs, gene regulatory networks; TF, transcription factor, TF_1, primary TF.

Upon nitrate exposure, NLP7 activates NRT2.1 expression (77). NLP7 also regulates NIGT1, a TF_2 that represses NRT2.1 activity. This coordination between the two TFs allows NRT2.1 expression to first rise and then fall within 3 h after nitrogen exposure (77).

In a coherent feed-forward loop, both the first TF [$TF_{2\text{-}1}$ (**Figure 4**)] and the second TF [$TF_{3\text{-}5}$ (**Figure 4**)] induce the target gene, and either one (OR logic) or both (AND logic) of the TFs are required for target expression (78). A coherent feed-forward loop with AND logic can act as a persistence detector, since short stimuli are insufficient for accumulation of the TF_2 (78). In contrast, a coherent feed-forward loop with OR logic shows a time delay in turning off the

response when the stimulus is removed, as the second TF maintains expression after the first TF is no longer active (3) (**Figure 4**). A series of coherent feed-forward loops regulate the dark-induced senescence pathway in which PIF4/PIF5 induce ORE1, an important TF involved in senescence, both directly and indirectly through EIN3 and ABI5/EEL (100). This process is thought to make the pathway robust and prevent transient fluctuations in light quality to influence gene regulation (100).

Another way TFs work together to generate dynamic gene expression patterns is through feedback loops. Feedback loops occur when a TF's expression is regulated by its downstream targets (90) (**Figure 4**). Feedback mechanisms can generate oscillations in gene expression and, thus, are useful when gene expression needs to be cyclical. Circadian rhythms, for example, are known to rely on feedback loops among multiple TFs in plants, animals, and fungi (59). In plants, LHY- and RVE8-clade TFs antagonistically regulate target genes and one another, creating several feedback loops to maintain correct functioning of the circadian clock across diverse environmental conditions (108). Many others plant processes rely on feedback mechanisms, such as transition to flowering (56) and maintenance of the stem cell niche (44).

GRNs can consist of interactions among tens to hundreds of individual TFs and thousands of targets with an intricate topology. Systems biology aims to simplify complex networks into more manageable components by identifying sets of TFs with similar properties and assessing how they are arranged into modules such as those described above. However, because of the many transient gene expression responses, short-lived TF–target interactions (see Sections 2 and 3), and a relatively limited quantity of genome-wide experimental data for temporal TF binding, most GRNs remain largely incomplete. In Section 5, we describe the computational approaches used to fill these gaps, largely by machine learning approaches that draw on known TF–target interactions and time-series gene expression data.

5. MODELING AND VALIDATING TIME-BASED GENE REGULATORY NETWORKS

Biological systems are incredibly dynamic. However, the current technologies available to probe the genome-wide interactions of regulators and their targets are largely static snapshots. This limitation means that most current GRN models are based on TF–target interaction data at one time point or a limited number of time points. A major goal of systems biology is to extend GRN models beyond such static measurements and learn networks, which will enable researchers to accurately predict how genes will respond in future, untested conditions.

The large quantity of genome-wide data being generated by next-generation sequencing, as described above (**Figure 1**; **Table 1**), makes machine learning a promising solution to the challenge of building GRNs able to predict future gene responses. This section aims to introduce the different computational techniques currently being used to infer the structure of GRNs and predict future expression levels of target genes, with a focus on time-based methods applied in plants. Since causality moves forward in time, time series represent a valuable data source for predictive networks. Thus, we focus on how time can be exploited in these predictive GRN models.

5.1. Coexpression Networks of Transcription Factors and Target Genes: Correlation Is Not Causality

Correlation has been widely adopted for the identification of TFs that regulate specific genes. Because of their low computational requirements and ease of use, correlation approaches to GRN construction have been extensively used in plant biology [see the review by Serin et al. (106)].

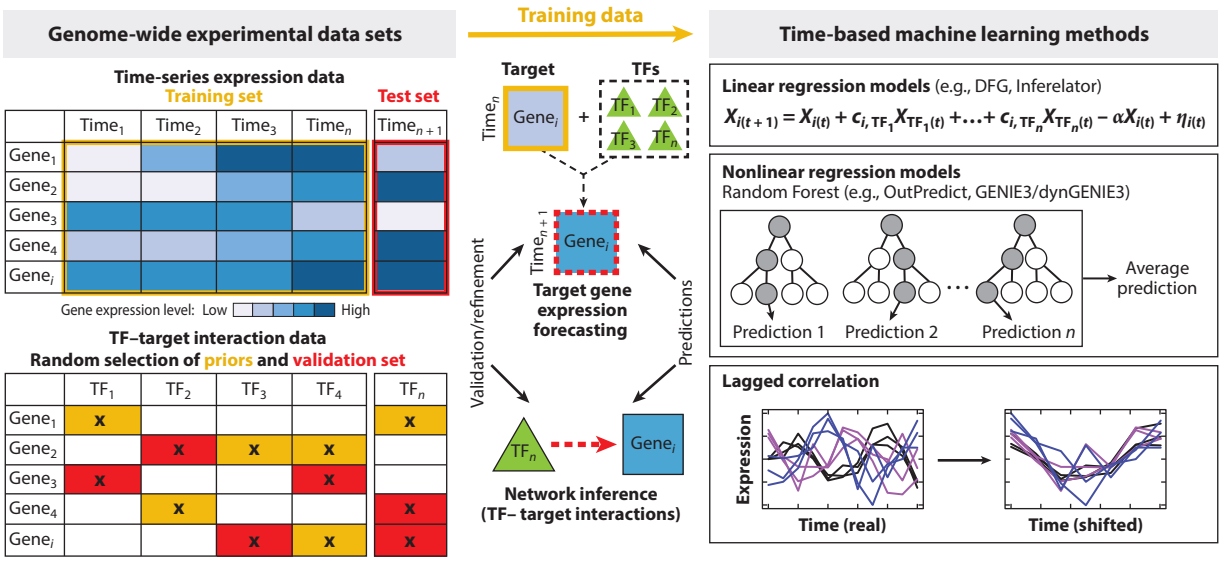

Figure 5

Systems biology methods to predict, validate, and refine GRNs from time-series gene expression data. Experimental data sets, such as genome-wide time-series transcriptomics and TF–target interactions, provide training data for time-based machine learning methods. The experimental data are often divided into a training/prior set, used to train the algorithm, and a test/validation set, used to validate and refine the resulting predicted GRN. A common machine learning method that uses time-series data is linear regression, in which the expression level (X) of Gene$_i$ at a given time point is predicted from the expression level of regulating TFs at the previous time points, each with a learned weight (c). Included in the model are a degradation term (a) and a noise term (η). Nonlinear regression approaches, such as methods using random forests, allow for more complex interactions between TFs and targets. These approaches predict causal interactions between TFs and targets, in contrast to correlation methods developed for time-series data, such as lagged correlation. While the most common goal is network inference, or the prediction of interactions between known TFs and target genes, another goal is target gene expression forecasting, or predicting the expression level of target genes at untested time points. The accuracy of predicted networks can be assessed by using the test/validation data sets. For time-series gene expression data, a so-called leave-out-last approach is often used. Here, the last time point (test set) is excluded from training, and predicted values are compared with actual (left-out) data to determine gene forecasting accuracy. For TF–target interactions, a randomly selected set of known, validated TF–target gene edges are excluded from the training data, and prediction accuracy can be tested using these left-out edges. This process can be repeated with different random selections of edges to determine the optimal algorithm parameters. Abbreviations: DFG, dynamic factor graph; GRN, gene regulatory network; TF, transcription factor.

However, one drawback to such correlation networks is that the interactions are undirected; that is, they lack information about causal relationships (93). Another disadvantage is that correlation networks do not necessarily distinguish between direct and indirect (i.e., downstream) TF–target interactions, potentially leading to many false-positive predicted interactions.

Time-series analysis of transcriptome data can be used in correlation network analysis to build GRNs (50). The delay between the transcription of a gene and the translation of the mRNA into protein means that a change in expression of a TF must precede that of its target genes. At high temporal resolution (e.g., minutes), this time delay can obscure interactions between TF regulators and their target genes in simple correlation networks. Thus, approaches that account for time lags in time-series data have been applied to correlation networks (60, 62, 104) (**Figure 5**; **Table 1**). Such time-lag correlation approaches reveal the temporal order within GRNs and allow directionality to be assigned to interactions between genes. However, they still cannot identify causal relationships. Additional data, such as known interactions from the literature, *cis*-binding

motif enrichment, and TF-binding assays (e.g., ChIP-seq), are therefore needed to provide additional support for interactions (10).

5.2. Time-Based Machine Learning Methods Can Infer Causality in Gene Networks and Forecast Target Gene Expression States

Forecasting: prediction of the future expression of target genes based on the expression of target gene and regulating TFs

DFG: Dynamic Factor Graph

Machine learning approaches are starting to improve in the areas where correlation analyses fall short, specifically, in predicting causal relationships between TFs and their target genes. A wide range of algorithms have been developed to address this challenge, and these various approaches and/or their mathematical frameworks have been reviewed elsewhere (9, 29, 40, 63, 64, 71, 101). Here, we focus on time-based machine learning methods that are being used specifically in plants to infer GRNs from time-series gene expression data.

As opposed to machine learning methods that use steady-state data, algorithms built for time-series data use the expression of the predictor variables at the previous time point to model the expression of target genes at the next time point. Its sequential nature makes time-series data particularly well suited to models that use regression analysis, which are able to describe the dynamic behavior of GRNs (62). The general approach of such analyses is to model the expression of a gene in the form of an ordinary differential equation, where it is a function of the expression of TFs (10, 54). Importantly, this approach allows the time-based models to perform not only network inference (i.e., predict interactions between influential TFs and their target genes) but also target gene forecasting (i.e., predict the future expression of target genes, at time $t + 1$, on the basis of the expression of the target gene and regulating TFs at the previous time point, time t). In addition, most time-based machine learning methods have some form of feature selection that delivers a sparse network that is more interpretable and better represents the scale-free structure of the biological network (86, 101, 124).

In linear regression, the coefficients for the TF predictors in the model provide a weight for the predicted interaction between the TF and the target gene (**Figure 5**; **Table 1**). An advantage of linear models is that they are mathematically simple and, therefore, often less computationally intensive (37). Alternatively, nonlinear regression models may better represent the actual biological system (101) (**Figure 5**; **Table 1**). Due to the flexibility and ease of use of these models, an increasingly popular approach to nonlinear regression is the use of random forest decision trees (53). One of the most popular random forest network inference algorithms is GENIE3 (54), but this method was built for steady-state data. GENIE3 has been adapted to take advantage of time-series data in order to infer the GRNs controlling the drought response in sunflower (82). More recently, a modification of GENIE3, dynamical GENIE3,was developed specifically to improve the handling of time-series data (41).

A dynamic factor graph (DFG) is an example of a linear regression approach that has been used to infer GRNs from fine-scale time-series transcriptomic data on the nitrogen response in *Arabidopsis* (20, 65, 128) (**Figure 5**; **Table 1**). In the first implementation of a DFG for GRN predictions (65), a fine-scale time-series experiment was performed and the ATH1 microarray was used to capture very early events in the nitrogen response. The DFG was then used to model the influence of 67 nitrogen-responsive TFs on one another, as well as gene targets in the nitrogen assimilation pathway. The authors of this study thus identified and functionally validated a novel TF regulator of nitrogen signaling, SPL9 (65). In subsequent time-series studies, a DFG was used to infer larger GRNs from fine-scale nitrogen-response time-course data generated using RNA-seq on shoots (of 172 TFs and 2,174 target genes) and roots (145 TFs and 1,458 target genes) (20, 128). These larger nitrogen-response GRNs were functionally validated by TF perturbation studies on a genome-wide scale, as discussed in Section 5.3, below.

ATAC-seq: assay for transposase-accessible chromatin using sequencing

The Inferelator is another popular algorithm that can generate linear models of gene expression from genome-wide data, including time series. In a study that employed the Inferelator to create GRNs for the response of rice cultivars to water and heat stress (130), known TF *cis*-binding motifs that were found in open chromatin regions, as determined by the use of assay for transposase-accessible chromatin using sequencing (ATAC-seq) on a subset of the experimental conditions, were used as priors to train the network models. This approach generated a final network connecting 113 TFs to 4,052 target genes (130).

The ability to improve network inference by incorporating priors has recently been demonstrated for a new random forest–based method called OutPredict (25). OutPredict is unique in that it allows the use of both steady-state and time-series data, as well as the incorporation of priors. It focuses on target gene expression forecasting, although it can also identify the underlying casual edges between a TF and a target gene. OutPredict has been applied to steady-state and time-series data sets from several different organisms, including *Arabidopsis*, and it outperformed several other inference methods, including dynamical GENIE3 (25).

Desai et al. (31) implemented a unique approach to time-based network inference that combined the expression level with the rate of change in expression, which they called ExRANGES. These authors showed that ExRANGES can improve network inference ability, compared with the use of expression values, when tested with both the Inferelator (15) and GENIE3 (54), which they modified to add a time-delay step (31). The authors suggested that this improvement is due to emphasizing time points that precede significant changes in target expression (31).

5.3. Addressing a Current Bottleneck in Gene Regulatory Networks: Genome-Wide Validation of Transcription Factor–Target Gene Predictions

Regardless of the machine learning approach used to predict TF–target gene relationships or forecasting target gene expression at untested time points, it is crucial to demonstrate that the TF–target predictions in the resulting GRNs are biologically relevant. Because a primary reason to build GRNs is to identify key regulators, highly connected TF hubs (see Section 4) are often selected for in planta validation (65, 128). However, this type of genome-wide validation on a single TF does not necessarily demonstrate the overall quality of the TF–target predictions generated for other TFs in the GRN. Therefore, it is also important to use statistical measures that will compare predictions of TF–target interactions in GRNs with actual values of experimentally validated TF–target interactions (105).

Beyond predicting TF–target gene interactions, time-series data can be used in machine learning models to forecast the expression level of a target gene at a future, untested time point. For methods that perform target gene expression forecasting, the models are often created from the gene expression data using a training set that includes all but the final time point, which is used as a test set (**Figure 5**). This approach, called leave-out-last, enables comparison between the accuracy of the predictions with the known (left-out) experimental values of target gene expression by different metrics. In one example of this approach, Krouk et al. (65) compared the sign of the change in target gene prediction with the actual sign of the change. They showed that the DFG method improved the predictions of target gene expression at the left-out time point by 20% in comparison to naive extrapolation of the gene expression based on the previous two time points (65). In another approach used in OutPredict, Cirrone et al. (25) compared the gene expression value returned for each prediction with the actual gene expression value in the test set and then calculated a global mean squared error metric. The predicted gene expression values from OutPredict were 34% more accurate than naive predictions for the *Arabidopsis* data set using the same metric.

In contrast, for network inference approaches, predicted regulatory edges between a TF and its target genes are compared with a gold standard network of known TF–target interactions (98). This comparison is usually done by calculating the number of true versus false positives and of true versus false negatives, which are then used to derive several metrics such as precision, recall, and specificity. When these metrics are taken at different cutoffs for the predicted network, they can be used to plot and calculate the area under the receiver operating characteristic curve and/or the area under the precision/recall (AUPR) curve, values used to compare the performance of inference methods to one another and to randomized networks (see References 98 and 105 for more details on these evaluation metrics). When gold standard edges are used as priors in network inference methods, they can be randomly split into a priors set and an evaluation set (**Figure 5**), allowing cross-validation of the predictions (105).

AUPR: area under the precision-recall

Time-series studies of the nitrogen response (20, 128) used precision/recall validation of TF–target interactions in the GRN to show that DFG-inferred networks generated from nitrogen by time-series transcriptome data have a better AUPR curve, and, thus, are able to predict true TF–target gene interactions significantly better than random networks (20, 128). Furthermore, these studies used the validated TF–target interactions for 7 TFs in the first study of shoots, followed by 33 TFs in the second study of roots, to refine the predicted network by selecting a precision cutoff value from the precision/recall curve and prune the inferred TF–target interactions to those that passed the cutoff TF–target edge value. This process allowed the authors to extend their validated data for ∼40 TFs to high-confidence predictions for more than 140 nitrogen-responsive TFs that regulate 720 nitrogen-responsive genes (20, 128).

Importantly, the gold standard interactions used to evaluate the networks in the above examples used validated genome-wide targets for 40 TFs in the nitrogen-response network (20, 128). A more recent analysis of nitrogen by the time DFG-inferred network showed that direct regulatory TF–target edges validated using root cell–based TARGET work better as a gold standard than TF–target interactions from in vitro binding assays (21). This result highlights the need for more TF perturbation data that validate TF–target regulation, which are currently more difficult to obtain and less scalable in comparison to in vitro TF-binding data.

There is a great deal of interest in developing new approaches, such as those described above, that are able to use time-series data to determine the structure of GRNs and predict gene expression. It is important to recognize that, while an individual method may outperform another using a specific data set, to date no algorithm that outperforms all other algorithms and is suitable for all types of data has been developed. In fact, ensemble approaches that combine the results of multiple inference methods can produce the most accurate predicted networks (81). Improving methods of handling time-series data is a key area that needs to be addressed, and to accomplish this goal it will be important to account for the dynamic nature of TF–target interactions, as described above.

6. PERSPECTIVE

Significant advances in addressing the challenge of capturing dynamic events in GRNs have been achieved. The methods described in this review have their own advantages and drawbacks but collectively can provide insight into the dynamic nature of GRNs.

How have the current advances changed our understanding of dynamic GRNs underlying plant responses to the environment? The discovery that TF–target interactions can occur in milliseconds and within 5 min of exposure to a given stimulus should thoroughly change our perspective of how rapidly regulatory events can take place. We propose that rapid and transient regulatory events may be more common than previously accepted. It is thus necessary to reevaluate the design of TF-binding experiments (such as ChIP) to capture TF dynamics. Plant cell–based

TF-perturbation assays such as TARGET (11, 20, 96) offer a higher-throughput experimental approach to the identification of these rapid and transient TF–target interactions (6, 96). However, such assays do not necessarily provide a native biological context, because they depend on cell isolation from tissues by protoplasting. In contrast, the dynamic and rapid changes in open chromatin regions provide clues to the effective regulatory regions relevant to transcriptional regulation in a native context. However, open chromatin studies fail to account for transient TF–target interactions and depend largely on the inference of TF-binding sites.

Studies of the function of multiple TFs in parallel have revealed the relative contribution of each TF and the TF hierarchy within a GRN. Importantly, TF–TF wiring is a prevalent component of TF networks that fine-tune gene expression responses to the environment. In particular, regulatory feedback and feed-forward loops are key motifs in dynamic GRNs.

A common limitation of the genome-wide studies discussed in this review is that they often mix different plant cell types together. Specifically, the vast majority of plant GRNs are derived from ChIP-seq or RNA-seq data sets where entire plant organs (e.g., root or shoot) are sequenced, in effect pooling cell types together. However, the handful of studies that have assayed cell type–specific expression patterns in plants have revealed that different cell types hold unique transcriptional profiles (13, 18) and respond to environmental signals in unique ways (33, 42). The recent advent of single-cell sequencing techniques offers an opportunity to create GRNs with cell type specificity. Already, single-cell profiling is able to correctly resolve distinct root cell types (30, 111, 132) as well as cell type differences in chromatin accessibility (36). Importantly, the differences between individual single-cell transcriptomes can be exploited to render a trajectory of cell type development in silico. Such "pseudotime" analyses can identify key genes involved in determining cell fate (111, 132) and likewise can be used to create dynamic GRNs (14). However, while such pseudotime analyses work for understanding cellular development, they are less helpful for modeling real-time changes in gene expression, especially in response to an environmental cue. Assaying plant environmental responses through single-cell sequencing is particularly difficult in plants in comparison to other organisms, as transcriptional responses can be compromised during the necessary step of digestion of plant cell walls (protoplasting). One solution to this problem is to perform single-nucleus sequencing instead, which would allow nuclei to be extracted from flash-frozen tissue, thus preserving transcriptomes within their native states (122).

Similarly, the development of machine learning approaches that can incorporate spatial information, in the form of single-cell or tissue-specific omics data, holds promise for improving predictions of dynamic GRNs, as was recently demonstrated for the GRNs controlling cell wall formation (114). Obtaining both spatial and temporal data will also enable exploration of how tissues and organs communicate. Making these connections between distinct populations of cells will require innovative computational approaches, such as the statistical approach recently used to connect nitrogen signals between the root and the shoot (49).

Single-cell multiomics technologies, assays for identifying TF–target gene binding and regulation, and machine learning methods are all rapidly evolving. Therefore, many challenges and opportunities remain on the horizon for mapping and modeling dynamic GRNs using plant systems biology.

SUMMARY POINTS

1. Time is an unexplored, underutilized, but powerful variable to interrogate the dynamics of GRNs and biological operating systems.

2. Single-molecule studies capture transcription factor (TF) binding dynamics within seconds but miss the genome-wide context. Conversely, genome-wide studies reveal the impacts of TF regulation genome-wide but often ignore the rapid and transient behavior of transcriptional regulation.
3. Chromatin immunoprecipitation sequencing (ChIP-seq) is the prevailing technique to capture TF-bound targets genome-wide. Time-series studies using ChIP-seq reveal TF dynamics. However, the TF binding captured by ChIP-seq rarely leads to TF regulation, and TFs are detectably bound to only a small proportion of their regulated targets.
4. Plant cell–based TF perturbation assays, such as TARGET (transient assay reporting on genome-wide effects of TFs), offer an experimental approach to the identification of rapid and transient genome-wide TF–target interactions.
5. The genome-wide relevance of multiple TFs can be captured in the same experimental setup by chromatin profiling assays such as DNAse-seq and assay for transposase-accessible chromatin using sequencing (ATAC-seq). Multiple TFs can also be studied in parallel by ChIP-seq or TARGET.
6. Network motifs, where two or more TFs work together, enable dynamic regulation of target gene expression.
7. Machine learning techniques are currently being exploited to predict causal relationships between TFs and their target genes.
8. The single-cell transcriptomics approach is a promising method to determine which GRNs drive dynamic responses in a high-resolution manner.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

Research on dynamic gene regulatory networks in G.C.'s laboratory is supported by National Institutes of Health (NIH) grant RO1 GM121753 and National Science Foundation Plant Genome Research Program grant IOS-1840761 to G.C., National Institute of General Medical Sciences Fellowship F32GM116347 to M.D.B., and Plant Genomics Grant A160051 from the Zegar Family Foundation. Research in J.M.A.'s laboratory is funded by Instituto Milenio iBio–Iniciativa Científica Milenio MINECON and Fondo Nacional de Desarrollo Científico y Tecnológico grant 1210389. J.S. is an Open Philanthropy awardee of the Life Sciences Research Foundation.

LITERATURE CITED

1. Akmakjian GZ, Bailey-Serres J. 2020. Nitrogen-responsive transcription factor kinetics meter plant growth. *PNAS* 117:13196–98
2. Albert R. 2005. Scale-free networks in cell biology. *J. Cell Sci.* 118:4947–57
3. Alon U. 2007. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* 8:450–61

4. Alvarez JM, Moyano TC, Zhang T, Gras DE, Herrera FJ, et al. 2019. Local changes in chromatin accessibility and transcriptional networks underlying the nitrate response in *Arabidopsis* roots. *Mol. Plant* 12:1545–60
5. Alvarez JM, Riveras E, Vidal EA, Gras DE, Contreras-López O, et al. 2014. Systems approach identifies TGA 1 and TGA 4 transcription factors as important regulatory components of the nitrate response of *Arabidopsis thaliana* roots. *Plant J.* 80:1–13

6. Demonstrates that NLP7 triggers a transcriptional cascade in response to nitrogen through transient TF–target interactions.

6. Alvarez JM, Schinke A-L, Brooks MD, Pasquino A, Leonelli L, et al. 2020. Transient genome-wide interactions of the master transcription factor NLP7 initiate a rapid nitrogen-response cascade. *Nat. Commun.* 11:1157
7. Aughey GN, Southall TD. 2016. Dam it's good! DamID profiling of protein-DNA interactions. *Wiley Interdiscip. Rev. Dev. Biol.* 5:25–37
8. Azpeitia E, Wagner A. 2020. Short residence times of DNA-bound transcription factors can reduce gene expression noise and increase the transmission of information in a gene regulation system. *Front. Mol. Biosci.* 7:67
9. Banf M, Rhee SY. 2017. Computational inference of gene regulatory networks: approaches, limitations and opportunities. *Biochim. Biophys. Acta Gene Regul. Mech.* 1860:41–52
10. Bar-Joseph Z, Gitter A, Simon I. 2012. Studying and modelling dynamic biological processes using time-series gene expression data. *Nat. Rev. Genet.* 13:552–64
11. Bargmann BO, Marshall-Colón A, Efroni I, Ruffel S, Birnbaum KD, et al. 2013. TARGET: a transient transformation system for genome-wide transcription factor target discovery. *Mol. Plant* 6:978–80
12. Becker M, Baumann C, John S, Walker DA, Vigneron M, et al. 2002. Dynamic behavior of transcription factors on a natural promoter in living cells. *EMBO Rep.* 3:1188–94
13. Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, et al. 2003. A gene expression map of the *Arabidopsis* root. *Science* 302:1956–60
14. Blencowe M, Arneson D, Ding J, Chen Y-W, Saleem Z, Yang X. 2019. Network modeling of single-cell omics data: challenges, opportunities, and progresses. *Emerg. Top. Life Sci.* 3:379–98
15. Bonneau R, Reiss DJ, Shannon P, Facciotti M, Hood L, et al. 2006. The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets de novo. *Genome Biol.* 7:R36
16. Bothma JP, Garcia HG, Esposito E, Schlissel G, Gregor T, Levine M. 2014. Dynamic regulation of eve stripe 2 expression reveals transcriptional bursts in living *Drosophila* embryos. *PNAS* 111:10598–603
17. Bothma JP, Norstad MR, Alamos S, Garcia HG. 2018. LlamaTags: a versatile tool to image transcription factor dynamics in live embryos. *Cell* 173:1810–22.e16
18. Brady SM, Orlando DA, Lee J-Y, Wang JY, Koch J, et al. 2007. A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318:801–6
19. Brewster RC, Weinert FM, Garcia HG, Song D, Rydenfelt M, Phillips R. 2014. The transcription factor titration effect dictates level of gene expression. *Cell* 156:1312–23
20. Brooks MD, Cirrone J, Pasquino AV, Alvarez JM, Swift J, et al. 2019. Network Walking charts transcriptional dynamics of nitrogen signaling by integrating validated and predicted genome-wide interactions. *Nat. Commun.* 10:1569

21. Constructs the ConnecTF platform to analyze publicly available data sets on TF–target gene interactions and validate predicted gene regulatory networks.

21. Brooks MD, Juang C-L, Katari MS, Alvarez JM, Pasquino A, et al. 2020. ConnecTF: a platform to integrate transcription factor–gene interactions and validate regulatory networks. *Plant Physiol.* In press. https://doi.org/10.1093/plphys/kiaa012
22. Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, et al. 2017. Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* 357:661–67

23. Characterizes EIN3 TF binding dynamics using ChIP-seq and addresses the contribution of EIN3 to transcriptional ethylene responses.

23. Chang KN, Zhong S, Weirauch MT, Hon G, Pelizzola M, et al. 2013. Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in *Arabidopsis*. *eLife* 2:e00675
24. Charoensawan V, Martinho C, Wigge PA. 2015. "Hit-and-run": Transcription factors get caught in the act. *BioEssays* 37:748–54
25. Cirrone J, Brooks MD, Bonneau R, Coruzzi GM, Shasha DE. 2020. OutPredict: Multiple datasets can improve prediction of expression and inference of causality. *Sci. Rep.* 10:6804
26. Clark NM, Hinde E, Winter CM, Fisher AP, Crosti G, et al. 2016. Tracking transcription factor mobility and interaction in *Arabidopsis* roots with fluorescence correlation spectroscopy. *eLife* 5:e14770

27. Concia L, Veluchamy A, Ramirez-Prado JS, Martin-Ramirez A, Huang Y, et al. 2020. Wheat chromatin architecture is organized in genome territories and transcription factories. *Genome Biol*. 21:104
28. Cuvier O, Fierz B. 2017. Dynamic chromatin technologies: from individual molecules to epigenomic regulation in cells. *Nat. Rev. Genet.* 18:457
29. Delgado FM, Gomez-Vela F. 2019. Computational methods for gene regulatory networks reconstruction and analysis: a review. *Artif. Intell. Med.* 95:133–45
30. Denyer T, Ma X, Klesen S, Scacchi E, Nieselt K, Timmermans MC. 2019. Spatiotemporal developmental trajectories in the *Arabidopsis* root revealed using high-throughput single-cell RNA sequencing. *Dev. Cell* 48:840–52.e5
31. Desai JS, Sartor RC, Lawas LM, Jagadish SVK, Doherty CJ. 2017. Improving gene regulatory network inference by incorporating rates of transcriptional changes. *Sci. Rep.* 7:17244
32. Ding Y, Fromm M, Avramova Z. 2012. Multiple exposures to drought 'train' transcriptional responses in *Arabidopsis*. *Nat. Commun.* 3:740
33. Dinneny JR, Long TA, Wang JY, Jung JW, Mace D, et al. 2008. Cell identity mediates the response of *Arabidopsis* roots to abiotic stress. *Science* 320:942–45
34. Doidy J, Li Y, Neymotin B, Edwards MB, Varala K, et al. 2016. "Hit-and-run" transcription: *De novo* transcription initiated by a transient bZIP1 "hit" persists after the "run." *BMC Genom*. 17:92
35. Dong P, Tu X, Chu P-Y, Lü P, Zhu N, et al. 2017. 3D chromatin architecture of large plant genomes determined by local A/B compartments. *Mol. Plant* 10:1497–509
36. Dorrity MW, Alexandre C, Hamm M, Vigil A-L, Fields S, et al. 2020. The regulatory landscape of *Arabidopsis thaliana* roots at single-cell resolution. bioRxiv 204792. **https://doi.org/10.1101/2020.07.17.204792**
37. Dubitzky W, Wolkenhauer O, Yokota H, Cho K-H, eds. 2013. *Encyclopedia of Systems Biology*. New York: Springer
38. Dufourt J, Trullo A, Hunter J, Fernandez C, Lazaro J, et al. 2018. Temporal control of gene expression by the pioneer factor Zelda through transient interactions in hubs. *Nat. Commun.* 9:5194
39. Eadara JK, Hadlock KG, Lutter LC. 1996. Chromatin structure and factor site occupancies in an in vivo–assembled transcription elongation complex. *Nucleic Acids Res.* 24:3887–95
40. Fogelberg C, Palade V. 2009. Machine learning and genetic regulatory networks: a review and a roadmap. In *Foundations of Computational Intelligence*, Vol. 1, ed. A-E Hassanien, A Abraham, AV Vasilakos, W Pedrycz, pp. 3–34. New York: Springer
41. Geurts P. 2018. dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data. *Sci. Rep.* 8:3384
42. Gifford ML, Dean A, Gutierrez RA, Coruzzi GM, Birnbaum KD. 2008. Cell-specific nitrogen responses mediate developmental plasticity. *PNAS* 105:803–08
43. Gitter A, Siegfried Z, Klutstein M, Fornes O, Oliva B, et al. 2009. Backup in gene regulatory networks explains differences between binding and knockout results. *Mol. Syst. Biol.* 5:276
44. Gordon SP, Chickarmane VS, Ohno C, Meyerowitz EM. 2009. Multiple feedback loops through cytokinin signaling control stem cell number within the *Arabidopsis* shoot meristem. *PNAS* 106:16529–34
45. Grönlund A, Lötstedt P, Elf J. 2013. Transcription factor binding kinetics constrain noise suppression via negative feedback. *Nat. Commun.* 4:1864
46. Gurdon JB, Javed K, Vodnala M, Garrett N. 2020. Long-term association of a transcription factor with its chromatin binding site can stabilize gene expression and cell fate commitment. *PNAS* 117:15075–84
47. Gutiérrez RA, Stokes TL, Thum K, Xu X, Obertello M, et al. 2008. Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. *PNAS* 105:4939–44
48. Hartmann L, Drewe-Boß P, Wießner T, Wagner G, Geue S, et al. 2016. Alternative splicing substantially diversifies the transcriptome during early photomorphogenesis and correlates with the energy availability in *Arabidopsis*. *Plant Cell* 28:2715–34
49. Heerah S, Molinari R, Guerrier S, Marshall-Colón A. 2020. Granger-causal testing for irregularly sampled time series with application to nitrogen signaling in *Arabidopsis*. bioRxiv 152819. **https://doi.org/10.1101/2020.06.15.152819**

30. Exploits single-cell transcriptomics to reveal key developmental regulators of root cell differentiation in Arabidopsis.

50. Hempel S, Koseska A, Nikoloski Z, Kurths J. 2011. Unraveling gene regulatory networks from time-resolved gene expression data—a measures comparison study. *BMC Bioinform*. 12:292
51. Hipp L, Beer J, Kuchler O, Reisser M, Sinske D, et al. 2019. Single-molecule imaging of the transcription factor SRF reveals prolonged chromatin-binding kinetics upon cell stimulation. *PNAS* 116:880–89
52. Hnisz D, Shrinivas K, Young RA, Chakraborty AK, Sharp PA. 2017. A phase separation model for transcriptional control. *Cell* 169:13–23
53. Huynh-Thu VA, Geurts P. 2019. Unsupervised gene network inference with decision trees and random forests. In *Gene Regulatory Networks: Methods and Protocols*, ed. G Sanguinetti, VA Huynh-Thu, pp. 195–215. New York: Springer
54. Huynh-Thu VA, Irrthum A, Wehenkel L, Geurts P. 2010. Inferring regulatory networks from expression data using tree-based methods. *PLOS ONE* 5:e12776
55. Inukai S, Kock KH, Bulyk ML. 2017. Transcription factor–DNA binding: beyond binding site motifs. *Curr. Opin. Genet. Dev.* 43:110–19
56. Jaeger KE, Pullen N, Lamzin S, Morris RJ, Wigge PA. 2013. Interlocking feedback loops govern the dynamic behavior of the floral transition in *Arabidopsis*. *Plant Cell* 25:820–33
57. Jariani A, Vermeersch L, Cerulus B, Perez-Samper G, Voordeckers K, et al. 2020. A new protocol for single-cell RNA-seq reveals stochastic gene expression during lag phase in budding yeast. *eLife* 9:e55320
58. Johnson KA, Goody RS. 2011. The original Michaelis constant: translation of the 1913 Michaelis–Menten paper. *Biochemistry* 50:8264–69
59. Jolma IW, Laerum OD, Lillo C, Ruoff P. 2010. Circadian oscillators in eukaryotes. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 2:533–49
60. Kiddle SJ, Windram OP, McHattie S, Mead A, Beynon J, et al. 2010. Temporal clustering by affinity propagation reveals transcriptional modules in *Arabidopsis thaliana*. *Bioinformatics* 26:355–62
61. Kimura H, Sugaya K, Cook PR. 2002. The transcription cycle of RNA polymerase II in living cells. *J. Cell Biol.* 159:777–82
62. Koryachko A, Matthiadis A, Ducoste JJ, Tuck J, Long TA, Williams C. 2015. Computational approaches to identify regulators of plant stress response using high-throughput gene expression data. *Curr. Plant Biol.* 3:20–29
63. Koryachko A, Matthiadis A, Muhammad D, Foret J, Brady SM, et al. 2015. Clustering and differential alignment algorithm: identification of early stage regulators in the *Arabidopsis thaliana* iron deficiency response. *PLOS ONE* 10:e0136591
64. Krouk G, Lingeman J, Marshall-Colón A, Coruzzi G, Shasha D. 2013. Gene regulatory networks in plants: learning causality from time and perturbation. *Genome Biol*. 14:123
65. Krouk G, Mirowski P, LeCun Y, Shasha DE, Coruzzi GM. 2010. Predictive network modeling of the high-resolution dynamic plant transcriptome in response to nitrate. *Genome Biol*. 11:R123
66. Kumar N, Singh A, Kulkarni RV. 2015. Transcriptional bursting in gene expression: analytical results for general stochastic models. *PLOS Comput. Biol.* 11:e1004292
67. Kuo MH, Allis CD. 1999. In vivo cross-linking and immunoprecipitation for studying dynamic protein:DNA associations in a chromatin environment. *Methods* 19:425–33
68. Lai X, Verhage L, Hugouvieux V, Zubieta C. 2018. Pioneer factors in animals and plants—colonizing chromatin for gene regulation. *Molecules* 23:1914
69. Lämke J, Brzezinka K, Altmann S, Bäurle I. 2016. A hit-and-run heat shock factor governs sustained histone methylation and transcriptional stress memory. *EMBO J.* 35:162–75
70. Larson DR, Zenklusen D, Wu B, Chao JA, Singer RH. 2011. Real-time observation of transcription initiation and elongation on an endogenous yeast gene. *Science* 332:475–78
71. Li Y, Pearl SA, Jackson SA. 2015. Gene networks in plant biology: approaches in reconstruction and analysis. *Trends Plant Sci*. 20:664–75
72. Li Y, Varala K, Coruzzi GM. 2015. From milliseconds to lifetimes: tracking the dynamic behavior of transcription factors in gene networks. *Trends Genetics* 31:509–15
73. Lickwar CR, Mueller F, Hanlon SE, McNally JG, Lieb JD. 2012. Genome-wide protein–DNA binding dynamics suggest a molecular clutch for transcription factor function. *Nature* 484:251–55
74. Liu K-H, Niu Y, Konishi M, Wu Y, Du H, et al. 2017. Discovery of nitrate–CPK–NLP signalling in central nutrient–growth networks. *Nature* 545:311–16

75. Loffreda A, Jacchetti E, Antunes S, Rainone P, Daniele T, et al. 2017. Live-cell p53 single-molecule binding is modulated by C-terminal acetylation and correlates with transcriptional activity. *Nat. Commun.* 8:313
76. MacQuarrie KL, Fong AP, Morse RH, Tapscott SJ. 2011. Genome-wide transcription factor binding: beyond direct target regulation. *Trends Genet.* 27:141–48
77. Maeda Y, Konishi M, Kiba T, Sakuraba Y, Sawaki N, et al. 2018. A NIGT1-centred transcriptional cascade regulates nitrate signalling and incorporates phosphorus starvation signals in *Arabidopsis*. *Nat. Commun.* 9:1376
78. Mangan S, Alon U. 2003. Structure and function of the feed-forward loop network motif. *PNAS* 100:11980–85
79. Mangan S, Zaslaver A, Alon U. 2003. The coherent feedforward loop serves as a sign-sensitive delay element in transcription networks. *J. Mol. Biol.* 334:197–204
80. Marand AP, Chen Z, Gallavotti A, Schmitz RJ. 2020. A *cis*-regulatory atlas in maize at single-cell resolution. bioRxiv 315499. **https://doi.org/10.1101/2020.09.27.315499**
81. Marbach D, Costello JC, Küffner R, Vega NM, Prill RJ, et al. 2012. Wisdom of crowds for robust gene network inference. *Nat. Methods* 9:796–804
82. Marchand G, Huynh-Thu VA, Kane NC, Arribat S, Varès D, et al. 2014. Bridging physiological and evolutionary time-scales in a gene regulatory network. *New Phytol.* 203:685–96
83. Marchive C, Roudier F, Castaings L, Bréhaut V, Blondet E, et al. 2013. Nuclear retention of the transcription factor NLP7 orchestrates the early response to nitrate in plants. *Nat. Commun.* 4:1713
84. McNally JG, Müller WG, Walker D, Wolford R, Hager GL. 2000. The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* 287:1262–65
85. Michaelis L, Menten ML. 1913. Die Kinetik der Invertinwirkung. *Biochem. Z.* 49:333–369
86. Michailidis G, d'Alché-Buc F. 2013. Autoregressive models for gene regulatory network inference: sparsity, stability and causality issues. *Math. Biosci.* 246:326–34
87. Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. 2002. Network motifs: simple building blocks of complex networks. *Science* 298:824–27
88. Mir M, Stadler MR, Ortiz SA, Hannon CE, Harrison MM, et al. 2018. Dynamic multifactor hubs interact transiently with sites of active transcription in *Drosophila* embryos. *eLife* 7:e40497
89. Morisaki T, Müller WG, Golob N, Mazza D, McNally JG. 2014. Single-molecule analysis of transcription factor binding at transcription sites in live cells. *Nat. Commun.* 5:4456
90. Muhammad D, Schmittling S, Williams C, Long TA. 2017. More than meets the eye: emergent properties of transcription factors networks in *Arabidopsis*. *Biochim. Biophys. Acta Gene Regul. Mech.* 1860:64–74
91. Neph S, Vierstra J, Stergachis AB, Reynolds AP, Haugen E, et al. 2012. An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* 489:83–90
92. O'Malley RC, Huang S-SC, Song L, Lewsey MG, Bartlett A, et al. 2016. Cistrome and epicistrome features shape the regulatory DNA landscape. *Cell* 165:1280–92
93. Opgen-Rhein R, Strimmer K. 2007. From correlation to causation networks: a simple approximate learning algorithm and its application to high-dimensional plant gene expression data. *BMC Syst. Biol.* 1:37
94. Paakinaho V, Presman DM, Ball DA, Johnson TA, Schiltz RL, et al. 2017. Single-molecule analysis of steroid receptor and cofactor action in living cells. *Nat. Commun.* 8:15896
95. Para A, Li Y, Coruzzi GM. 2018. μChIP-seq for genome-wide mapping of in vivo TF-DNA interactions in *Arabidopsis* root protoplasts. In *Root Development*, ed. D Ristova, E Barbez, pp. 249–61. New York: Springer
96. Para A, Li Y, Marshall-Colón A, Varala K, Francoeur NJ, et al. 2014. Hit-and-run transcriptional control by bZIP1 mediates rapid nutrient signaling in *Arabidopsis*. *PNAS* 111:10371–76
97. Pfeiffer A, Shi H, Tepperman JM, Zhang Y, Quail PH. 2014. Combinatorial complexity in a transcriptionally centered signaling hub in *Arabidopsis*. *Mol. Plant* 7:1598–618
98. Prill RJ, Marbach D, Saez-Rodriguez J, Sorger PK, Alexopoulos LG, et al. 2010. Towards a rigorous assessment of systems biology models: the DREAM3 challenges. *PLOS ONE* 5:e9202
99. Rennie S, Dalby M, Van Duin L, Andersson R. 2018. Transcriptional decomposition reveals active chromatin architectures and cell specific regulatory interactions. *Nat. Commun.* 9:487

88. Discovers that Zelda and Bicoid interact only transiently with sites of active transcription in *Drosophila* embryos.

92. Develops a DAP-seq approach that identifies genomic regions bound by TFs in vitro and reveals that TF binding is DNA methylation sensitive.

100. Sakuraba Y, Jeong J, Kang M-Y, Kim J, Paek N-C, Choi G. 2014. Phytochrome-interacting transcription factors PIF4 and PIF5 induce leaf senescence in *Arabidopsis*. *Nat. Commun.* 5:4636
101. Sanguinetti G. 2019. Gene regulatory network inference: an introductory survey. In *Gene Regulatory Networks: Methods and Protocols*, ed. G Sanguinetti, VA Huynh-Thu, pp. 1–23. New York: Springer
102. Sayou C, Nanao MH, Jamin M, Posé D, Thévenon E, et al. 2016. A SAM oligomerization domain shapes the genomic binding landscape of the LEAFY transcription factor. *Nat. Commun.* 7:11222
103. Schaffner W. 1988. A hit-and-run mechanism for transcriptional activation? *Nature* 336:427–28
104. Schmitt WA, Raab RM, Stephanopoulos G. 2004. Elucidation of gene interaction networks through time-lagged correlation analysis of transcriptional data. *Genome Res.* 14:1654–63
105. Schrynemackers M, Küffner R, Geurts P. 2013. On protocols and measures for the validation of supervised methods for the inference of biological networks. *Front. Genet.* 4:262
106. Serin EA, Nijveen H, Hilhorst HW, Ligterink W. 2016. Learning from co-expression networks: possibilities and challenges. *Front. Plant Sci.* 7:444
107. Shah M, Funnell AP, Quinlan KG, Crossley M. 2019. Hit and run transcriptional repressors are difficult to catch in the act. *BioEssays* 41:1900041

108. Reveals that multiple feedback loops of the plant clock help ensure rhythmicity under adverse environmental conditions.

108. Shalit-Kaneh A, Kumimoto RW, Filkov V, Harmer SL. 2018. Multiple feedback loops of the *Arabidopsis* circadian clock provide rhythmic robustness across environmental conditions. *PNAS* 115:7147–52

109. Sharma R, Singh G, Bhattacharya S, Singh A. 2018. Comparative transcriptome meta-analysis of *Arabidopsis thaliana* under drought and cold stress. *PLOS ONE* 13:e0203266
110. Shimbo T, Du Y, Grimm SA, Dhasarathy A, Mav D, et al. 2013. MBD3 localizes at promoters, gene bodies and enhancers of active genes. *PLOS Genet.* 9:e1004028
111. Shulse CN, Cole BJ, Ciobanu D, Lin J, Yoshinaga Y, et al. 2019. High-throughput single-cell transcriptome profiling of plant cell types. *Cell Rep.* 27:2241–47.e4
112. Sigova AA, Abraham BJ, Ji X, Molinie B, Hannett NM, et al. 2015. Transcription factor trapping by RNA in gene regulatory elements. *Science* 350:978–81

113. Combines the binding profiles of 21 TFs with differentially expressed genes in response to ABA to reveal a stress response network.

113. Song L, Huang SC, Wise A, Castanon R, Nery JR, et al. 2016. A transcription factor hierarchy defines an environmental stress response network. *Science* 354:aag1550

114. Develops a machine learning approach that integrates expression, TF–DNA binding, and open chromatin data to infer condition-specific regulatory genes.

114. Song Q, Lee J, Akter S, Rogers M, Grene R, Li S. 2020. Prediction of condition-specific regulatory genes using machine learning. *Nucleic Acids Res.* 48:e62

115. Spivakov M. 2014. Spurious transcription factor binding: non-functional or genetically redundant? *BioEssays* 36:798–806
116. Sullivan AM, Arsovski AA, Lempe J, Bubb KL, Weirauch MT, et al. 2014. Mapping and dynamics of regulatory DNA and transcription factor networks in *A. thaliana*. *Cell Rep.* 8:2015–30
117. Swift J, Alvarez JM, Araus V, Gutiérrez RA, Coruzzi GM. 2020. Nutrient dose-responsive transcriptome changes driven by Michaelis–Menten kinetics underlie plant growth rates. *PNAS* 117:12531–40
118. Swift J, Coruzzi GM. 2017. A matter of time—how transient transcription factor interactions create dynamic gene regulatory networks. *Biochim. Biophys. Acta Gene Regul. Mech.* 1860:75–83
119. Szabo Q, Bantignies F, Cavalli G. 2019. Principles of genome folding into topologically associating domains. *Sci. Adv.* 5:eaaw1668
120. Tao Z, Shen L, Gu X, Wang Y, Yu H, He Y. 2017. Embryonic epigenetic reprogramming by a pioneer transcription factor in plants. *Nature* 551:124–28
121. Tanay A. 2006. Extensive low-affinity transcriptional interactions in the yeast genome. *Genome Res.* 16:962–72
122. Tian C, Du Q, Xu M, Du F, Jiao Y 2020. Single-nucleus RNA-seq resolves spatiotemporal developmental trajectories in the tomato shoot apex. bioRxiv 305029. **https://doi.org/10.1101/2020.09.20.305029**
123. Tian F, Yang D-C, Meng Y-Q, Jin J, Gao G. 2020. PlantRegMap: charting functional regulatory maps in plants. *Nucleic Acids Res.* 48:D1104–13
124. Tibshirani R. 1996. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* 58:267–88
125. Todeschini A-L, Georges A, Veitia RA. 2014. Transcription factors: specific DNA binding and specific gene regulation. *Trends Genet.* 30:211–19

126. Torti S, Fornara F, Vincent C, Andrés F, Nordström K, et al. 2012. Analysis of the Arabidopsis shoot meristem transcriptome during floral transition identifies distinct regulatory patterns and a leucine-rich repeat protein that promotes flowering. *Plant Cell* 24:444–62
127. Varala K, Li Y, Marshall-Colón A, Para A, Coruzzi GM. 2015. "Hit-and-run" leaves its mark: catalyst transcription factors and chromatin modification. *BioEssays* 37:851–56
128. Varala K, Marshall-Colón A, Cirrone J, Brooks MD, Pasquino AV, et al. 2018. Temporal transcriptional logic of dynamic regulatory networks underlying nitrogen signaling and use in plants. *PNAS* 115:6494–99
129. Wang M, Wang P, Lin M, Ye Z, Li G, et al. 2018. Evolutionary dynamics of 3D genome architecture following polyploidization in cotton. *Nat. Plants* 4:90
130. Wilkins O, Hafemeister C, Plessis A, Holloway-Phillips M-M, Pham GM, et al. 2016. EGRINs (environmental gene regulatory influence networks) in rice that function in the response to water deficit, high temperature, and agricultural environments. *Plant Cell* 28:2365–84
131. Zander M, Lewsey MG, Clark NM, Yin L, Bartlett A, et al. 2020. Integrated multi-omics framework of the plant response to jasmonic acid. *Nat. Plants* 6:290–302
132. Zhang T-Q, Xu Z-G, Shang G-D, Wang J-W. 2019. A single-cell RNA sequencing profiles the developmental landscape of *Arabidopsis* root. *Mol. Plant* 12:648–60
133. Zhang W, Zhang T, Wu Y, Jiang J. 2012. Genome-wide identification of regulatory DNA elements and protein-binding footprints using signatures of open chromatin in Arabidopsis. *Plant Cell* 24:2719–31

128. Exploits time-series data and machine learning algorithms to uncover gene regulatory networks controlling dynamic responses to nitrogen.