Temporal correlations robustly reveal regulatory coherence upon environmental perturbation

Haoran Cai[1], David L. Des Marais[1]

[1]Department of Civil and Environmental Engineering, MIT.
15 Vassar St., Cambridge, MA, 02139 USA

KEYWORDS: Network rewiring; Context-specific network; Temporal transcriptome; Dynamic correlation; Regulatory coherence

Abstract

Transcriptional Regulatory Networks (TRNs) orchestrate the timing, magnitude, and rate of organismal response to many environmental perturbations. Regulatory interactions in TRNs are context-dependent, affording the opportunity to detect responsive links upon environmental perturbation. To measure the transition between regulatory status directly, one powerful approach is chromatin immuno-precipitation binding, ChIP-Seq. However, genome-wide ChIP-Seq is costly and currently inaccessible for many transcription factors (TFs) across multiple conditions in most organisms. Seeking to exploit the abundance of RNASequencing data now available, many past studies have relied upon population-level statistics from cross-sectional studies, generating pairwise gene co-expression, to capture transient regulatory activity. Here, we employ a minimal stochastic model of transcriptional regulation and demonstrate that population correlations from cross-sectional studies may fail to capture transient regulatory activities. To characterize network rewiring in response to environmental perturbations, we use the dynamic correlation between time-series gene expression profiles in *Oryza sativa* in a network prior. Overall, stronger regulatory interactions are observed following environmental perturbation, which we term regulatory coherence. Previously known regulators with changing regulatory activities are reliably identified with our method. Exploiting dynamic correlations has the potential to prioritize stress-responsive regulators, affording greater detection power as compared to traditional differential expression approaches.

1 Introduction

Transcriptional regulation plays a fundamental role in the behavior of cells responding to external and internal environmental cues, and is aberrant in many diseases (Lee and Young 2013). Networks comprised of directed links between pairs of genes are termed "Transcriptional regulatory networks" (TRNs), and are often used by organisms to coordinate cellular, physiological and developmental response to the varying external environment. These "plastic" responses are believed to be particularly important for plants which, as sessile organisms, must adjust their physiology and development to cope with variable environments (Van Kleunen and Fischer 2005). TRNs orchestrate the timing and rate of genome-wide gene expression plastic responses (Gibson 2008). A long-standing goal in molecular biology is to understand functional linkages between regulatory architectures in TRNs and the dynamic behavior of organisms (Smith 1990). Transcription factors (TFs) are key nodes in TRNs that regulates the expression of other genes

(Buchanan et al. 2010), coordinating the entire transcriptional program. Accordingly, TFs remain appealing therapeutic and engineering targets for disease (Neef et al. 2011, Spampanato et al. 2013, Neef et al. 2010, Cuadrado et al. 2018) and stress tolerance for crop production (Wang et al. 2016, Lan Thi Hoang et al. 2017). As large-scale high-throughput -omic data are now readily generated for many ecologically and economically important species (Chen et al. 2020, Li et al. 2019, Müller et al. 2017, Rastogi et al. 2018, Wilkins et al. 2016), data-driven methods are needed for discovering key responsive TFs and for understanding how such TFs drive dynamic responses to the environment.

Many regulatory interactions in TRNs are context-dependent (Dunlop et al. 2008, Luscombe et al. 2004). During the lifetime of an organism, TRNs are dynamic in response of environmental cues and the actual transcriptional regulations can change. Environmental perturbations often trigger structured gene expression response (Bar-Joseph et al. 2012) and, by extension, interactions between transcripts. For instance, a protein that is an activator in one context may act as a repressor in another (Chahrour et al. 2008, Campbell et al. 2004). Moreover, an interaction can be inactive simply because the concentration of the regulator is outside its effective range for the target (Dunlop et al. 2008), or as a result of the chromatin state of target genes, post-translational modification of the regulator itself, the presence of inhibitory factors, or the absence of co-factors (Toledo and Wahl 2006, Piggot and Hilbert 2004). Identifying transient status changes for regulatory interactions upon environmental perturbation is critical to understand how genes and proteins modulate cellular and organismal responses to variable environments. One powerful approach to directly measure regulatory interactions is to use ChIP-seq. However, genome-wide ChIP-seq for hundreds of transcription factors across multiple conditions is costly, technically challenging, and therefore currently inaccessible in most organisms. Previous studies in environmental responses using ChIP-seq have focused on a small set of TFs or in a specific environmental condition (Chang et al. 2013, Ernst et al. 2007, Harbison et al. 2004, Hahn et al. 2004). Temporal gene expression datasets may capture transient response in perturbation-response experiments, during which genes may respond asynchronously with complicated kinetics (Gasch et al. 2000, Bar-Joseph et al. 2012). Therefore, extensive efforts have been put to exploit the so-called "fourth dimension" of response — time — to better understand the dynamics of TRNs and to identify putative signaling pathways or key regulatory genes (Bechtold et al. 2016, Yeung et al. 2018, Varala et al. 2018, Zander et al. 2020, Song et al. 2016, Greenham et al. 2017, Windram et al. 2012, Gargouri et al. 2015, Alvarez-Fernandez et al. 2020). However, few of these earlier studies examined context-dependent regulatory interactions. Here we develop an

approach to infer context-dependent regulatory status using a combination of ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) and temporal gene expression data.

The context-dependency of many regulatory interactions affords the opportunity to utilize the change of correlation between transcripts as a means to prioritize genes for further study (Luscombe et al. 2004, Duren et al. 2020). We hypothesize that responsive interactions under a given stimulus will alter the regulatory status transiently, reflected by a higher coexpression between transcript abundances. At genome scale, such increased correlations among genes in response to environmental perturbations may be termed regulatory coherence. Conversely, a recent study found that internal and external stressors can cause regulatory "decoherence" (lower correlation) in broad contexts (Lea et al. 2019). Here we use the dynamic correlation between time series gene expression profiles in *Oryza sativa* to characterize the contrasting status of regulatory interactions upon environmental perturbations — coherence vs decoherence — which may reflect the transient dynamic of TF-target binding. Our goal is not to infer a static TRN from transcriptome-wide gene expression data but, rather, to reconstruct a dynamic network under different conditions to detect environmentally induced links. Given the fact that gene regulation occurs with time-delay (Munsky et al. 2012), our measure takes into account the staggered relationships between the change of expression of a regulator and its putative target (Fig. S1C). This staggered relationship reflects the time required for a given TF activity to influence the expression of target genes due to the fact that transcription and translation take place over non-negligible time periods. To better measure the regulatory strength and characterize context-specific transcriptional network, we propose to maximize the cross-correlation over a time course in a directed interaction (from regulator to target) to evaluate its regulatory status. We investigated whether external stresses cause overall regulatory coherence in the rice TRN by applying our method to study time courses of heat- and drought- induced gene expression. We explored the determinants of differential regulatory strength and found that responsive TFs and links appear to drive regulatory coherence. In addition, we demonstrate that our time course regulatory analysis pipeline can perform better than the traditional differential expression approach in detecting responsive TFs, reliably dissecting the dynamics at a network level.

2 Results

2.1 Environmental stress leads to transient regulatory coherence

The expression relationship observed between genes in a time series sample may be caused by the time-lag inherent in molecular interactions, in this case, transcriptional regulation. Such time lag reflects the time required for a TF's activity to influence the expression of its target genes because transcription and translation take place over non-negligible time periods. Here we used a metric we call the regulatory score, building on the cross-correlation between transcript abundances estimated by RNA-Seq, to examine the differential regulatory interactions in a rice TRN obtained from Wilkins et al. (2016). The essential idea of the algorithm is to find a peak correlation between sampling points along a time series with time-lags (Fig. S1A - D). As such, our method provides a straightforward statistical means to generate hypotheses about TF-gene interactions using the now widely abundant RNA-Sequencing data.

We calculated the regulatory scores for each pair in the network prior per cultivar per condition with a time duration of 135 minutes immediately after the incidence of stress (elevated temperature or dehydration). Calculated regulatory scores from all four cultivars were merged stress-wise to increase sample size. We set a threshold of 0.54 (FDR = 5%) according to the regulatory score under controlled conditions as the cutoff of activation for regulatory interaction. We observed evidence for overall regulatory coherence. Here, we use the terms regulatory coherence and decoherence to mean increasing or decreasing co-expression, respectively. Coherence in our analyses is reflected by fewer time-lag τ_{reg} and higher regulatory scores under the heat and dehydration stress conditions when compared to control samples, as imposed by Wilkins et al. (2016). First, both drying and heat stresses result in more synchronized gene expression with decreased τ_{reg} (p value $<$ 2.2e-16 for both heat and drought stress, binomial test, Fig. S2) . Second, both stressors lead to higher regulatory scores for regulator-target pairs in the network prior.

Overall, the distribution of regulatory scores (Figure S1E) reveals that stressful environments increase the regulatory strength among measured transcripts in the network prior. The distribution of the regulatory score under heat (Kolmogorov-Smirnov test statistic D = 0.0445, p-value $< 2.2 \times 10^{-16}$) and drought (Kolmogorov-Smirnov test statistic D = 0.0929, p-value $< 2.2 \times 10^{-16}$) conditions are significantly different from the control condition. We identified significant TF-gene interactions in stressful conditions which were not observed in the controlled condition, and vice versa. We found greater support for the former number: out of 38127 total

interactions in the network prior, 3805 and 5965 pairs turn into active pairs from inactive pairs in heat and drying stress, respectively (Fig. 1), whereas only 1472 and 1604 of them possess opposite patterns under heat or drying, respectively. The observation of regulatory coherence is robust to various thresholds for activation (Fig. S6 and Fig. S7). Collectively, these results suggest a strong bias toward regulatory coherence in the rice data.

2.2 Population-level correlation can lead to temporal bias in detecting environmentally-responsive regulatory links

Our observation that co-expression increases (regulatory coherence) with the onset of stress is seemingly inconsistent with a recent study which used gene expression profiles collected from human monocytes exposed to a chemical stress in vitro to calculate differential population correlation among transcripts (Lea et al. 2019). Lea et al. found evidence supporting regulatory decoherence in broad contexts. We here posit that population-level differential correlations (cross-sectional correlation) at two time points may confound observations of regulatory coherence. Here, through a minimal mathematical model, we illustrate how saturation in gene regulatory interactions at the molecular level may be a confounding factor when using population-level correlation, leading to potential bias.

We contrast the performance of population-level metric with our measure based on cross-correlation. A typical regulation function between a TF and a target gene (modeled as a dose-response curve) can be characterized as Hill function, which is a nonlinear function (cooperative binding mode with Hill coefficient $n > 1$) as shown in Fig. 2C and F (grey line). Two perturbation regimes are considered: A saturated regime in which additional TF transcripts beyond some concentration threshold fail to induce additive response in their target genes, and a non-saturated regime characterized as the portion of a dose-response curve in which additional TF transcripts are associated with increased expression of their targets. We assume that the external perturbation modulates gene expression dynamics by the signal S_x. Smaller K_x and larger Hill coefficients increase the probability of saturation of regulation after environmental perturbation. The saturation of the regulation effectively masks the differential regulatory interaction upon perturbation, even if the TF X is nominally an environmentally induced activator of the gene Y. In addition, two possible external perturbation regimes are simulated (Fig. 2B): Press perturbation and pulse perturbation, which differ in the duration of the perturbation imposed at the given regulatory pair. If the upstream signal for a TF-gene pair has the property of adaptation, the signal induction may only last for a short period of time. Adaptation here is

defined by the ability of circuits to respond to input change but to return to the pre-stimulus output level, even when the input change persists (Ma et al. 2009, Briat et al. 2016).

Fig. 2E shows that, under a saturated regime, the population-level correlation between X and Y becomes lower under a perturbation, despite the fact that the interaction between X and Y is activated by an environmental perturbation. On the other hand, under a non-saturated regime, the population-level correlation increases (Fig. 2H). It should be noted that under a non-cooperative binding mode (e.g., Hill coefficient equals 1), the population-level correlation will decrease independently under the saturation regime. Therefore, how population-level correlations change relies upon whether a given transcriptional regulation is under a saturated regime or not; population-level correlations can fail to capture transient environmentally responsive links. Such bias can be termed as the temporal bias, which has been found to undermine the predictive power of machine learning models in medicine (Yuan et al. 2021). However, the temporal correlation between TF X and target gene Y is sensitive enough to characterize the environmentally induced activation under both saturated and non-saturated regimes induced by either press or pulse perturbation by S_x (Fig. 2D and G). These results highlight the likely high incidence of false negatives in identifying responsive gene interactions when relying on population-level correlations. Furthermore, these results imply that correlation statistics derived from temporal dynamics may be better suited for differential network analysis and to robustly identify environmentally responsive links.

2.3 Prioritizing novel candidates in regulating stress responses

Based on our observation that regulatory coherence increases transiently under the two stressors studied herein, we next investigated what features of TRNs might determine regulatory coherence. First, we examined whether certain TF families affect the activation of regulatory interactions: Each regulator-target pair is assigned a family according to the regulator (Fig. 3A). We found that many TFs with high differential regulatory scores in the heat-stress data set are annotated as Heat Shock transcription Factors (HSFs). As a family, HSFs have been demonstrated previously to show an important role in regulating genome-wide responses to elevated temperature in diverse species (Wang et al. 2004, Ohama et al. 2017).

Inspecting the relationship between differential gene expression and the differential regulatory score reveals that several known HSFs that do not independently show strong expression response to the stressor do show a clear response according to the differential regulatory score. We also find several interesting candidate TFs outside of the HSF family which have high differential

regulatory scores but little or no apparent differential expression in pairwise contrasts between control and treatment conditions (Fig. 4A). In the heat data set, the TF OsTCP7 has a differential regulatory score of 0.54 but was not identified as differentially expressed by Wilkins et al. (Wilkins et al. 2016). TCPs are broadly involved in regulating cell proliferation and growth (Martín-Trillo and Cubas 2010) and so OsTCP7 may be an interesting candidate for functional validation in the context of heat stress response.

While the HSF TFs comprise a gene family and are generally interpreted as coordinating plant response to heat stress (von Koskull-Döring et al. 2007), the regulatory control of response to soil drying is more distributed among diverse gene families and regulatory pathways (Joshi et al. 2016, Manna et al. 2020). Our analysis of the drought response data identified several TFs with previously demonstrated roles in rice dehydration response. These include HOX24 (Bhattacharjee et al. 2021) and ZFP182 (Huang et al. 2012), both of which were also found by Wilkins et al. to show a strong differential response. Several interesting candidates emerge among the list of TFs which have high differential regulatory score but low differential expression response. One such gene is PIF-Like 12 (Nakamura et al. 2007) which, to our knowledge, has no known role in dehydration response but is paralagous to OsPIL1 which integrates cues from the circadian clock and dehydration signaling to control internode elongation in rice (Todaka et al. 2012). Additional candidate genes with high regulatory scores under elevated temperature or dehydration stress are shown in Table S1 and S2. We hypothesize that the differential regulatory score could be used to identify novel stress-responsive regulators.

3 Discussion

Studies in human disease and plant and animal stress response frequently use genome-wide gene expression data to study how co-expression changes and networks are rewired during environmental perturbation (Fukushima 2013, Southworth et al. 2009, de la Fuente 2010, Amar et al. 2013, Choi et al. 2005, Kostka and Spang 2004, Deng et al. 2015, Yan et al. 2019, Cho et al. 2009, Fukushima et al. 2012, de la Fuente 2010, Li and Li 2020). Ultimately, the aim of such studies is to identify the cellular basis of environmental responses as a means to identify abnormal regulation in disease state and to improve medical interventions (Southworth et al. 2009, de la Fuente 2010, Amar et al. 2013, Kostka and Spang 2004), design breeding strategies (Fukushima et al. 2012), or to parameterize models of molecular evolution (Wray et al. 2003). However, many past studies have relied on population-level statistics to generate pairwise gene co-expression to detect transiently changing interactions (Cortijo et al. 2020, Fukushima et al.

2012, Deng et al. 2015, Lea et al. 2019). While statistically straightforward, these widely used methods likely miss many interactions which drive organismal response and therefore provide an incomplete picture of these complex systems. First, without a static transcriptional network prior, generating pairwise gene co-expression network and detecting responsive links can lead to false positives since many of links may be indirect and not involve any causal regulatory relationship (Feizi et al. 2013, Barzel and Barabási 2013). Second, population-level statistics are often confounded by individual covariates such as genotype, age, and sex (Parsana et al. 2019, Lea et al. 2019). Third, even within an isogenic homogeneous population (e.g., single snapshot single-cell RNA-seq collected under control versus perturbed condition), cross sectional population correlations may be confounded by switch-like transitions and ultrasensitivity in gene regulation, thereby failing to characterize the dynamic network rewiring (See Fig. 2).

In the present study, we implemented a stochastic simulation of a minimal regulation model under two perturbation regimes (Fig. 2C and 2F) to examine the performance in detecting responsive links with either temporal or population-level data. We show that, under a cooperative binding mode, the population-level co-expression changes of an environmentally induced link depend upon whether the gene regulation is under the saturated regime or not. Our results also indicate that while population-level correlations may be confounded by saturated regulation, dynamic correlations derived from temporal gene expression are robust in both regimes. Hence, we reconcile seemingly inconsistent observations with Lea et al. (2019): While co-expression tends to be transiently coherent, whether the co-expression measured using population statistics becomes coherent or decoherent depends upon the specific parameters of a given gene pair and the environmental condition. We term such bias that occurs when using cross-sectional data (herein population-level correlation) as temporal bias. As such, we maintain that to capture transient changes of the regulatory interaction and reconstruct context-specific transcriptional networks, longitudinal data is better than two time point cross-sectional data.

3.1 Limits of single-cell snapshots

Our work sheds light on the increasing prevalent data from scRNA-sequencing in studying perturbation response (Azizi et al. 2018, Li and Li 2020). Genome-scale high-dimensional measurements on single cells have been readily generated recently, which can be applied at relatively low cost to thousands and even tens of thousands of cells. An ambitious goal of single-cell analysis is to describe dynamic cell behavior and progressions of biological processes and, by extension, to reveal dynamic gene regulation. Our result illustrates that, in perturbation-

response studies, the reconstruction of the dynamic network maybe confounded by extracting information from static snapshots alone (time-stamped cross-sectional data), thereby leading to false negative environmentally responsive links. Therefore, our result emphasizes the need to reconstruct the temporal dynamics indirectly from population-level snapshots. Indeed, a number of methods (Specht and Li 2017, Matsumoto et al. 2017, Campbell and Yau 2016, Reid and Wernisch 2016, Äijö et al. 2014, Campbell and Yau 2015) have been trying to infer the latent temporal progression from cross sectional transcriptomic data and have been used to order events in cell differentiation (Bendall et al. 2014, Moignard et al. 2015), cancer progression (Sun et al. 2021) and cell cycle regulation (Kafri et al. 2013).

3.2 Coherence or decoherence?

Our observation that co-expression increases with the onset of stress is seemingly inconsistent with a recent study which used gene expression data collected from human monocytes to infer population correlation among transcripts (Lea et al. 2019). We established that the rationale underlying the inconsistent results could be that the population-level correlation can fail to detect transient coherence when contrasting pools of stressed vs unstressed samples, arising from the fact that statistically detectable patterns of co-variance may be transient (Fig 2). Further analysis is needed to examine to what extent this type of saturated regulatory regime is pervasive. Beyond this technical explanation, biological phenomena may underlie the seemingly inconsistent observation. First, our results suggest that the environmentally activated regulatory interactions become more temporally coherent in the presence of external stimuli whereas Lea et al. examined differential co-expression among all differentially expressed genes. It is possible that pairs without an actual regulatory relationship as well as non-responsive pairs with a regulatory relationship tend to be decoherent upon external stimuli. Therefore, without a network prior and therefore knowledge of an expected regulatory interaction, it would be more likely to observe regulatory decoherence. Second, the co-expression pattern itself may be strongly context dependent; the differential co-expression pattern likely depends on the severity and duration of stress. Linking the technical reasons above, the severity of stress can lead to higher probability of saturated regulatory regimes, hence higher probability of false negatives when using population-level correlation. Third, the sampling time may be critical. We examined the immediate response and transient dynamics of the regulation while Lea et al. examined the correlation patterns two hours after the onset of an external stimulus.

From an evolutionary perspective, a commonly observed result is that phenotypic integration in a population (i.e., the number of significant correlations among traits) increases with environmental stress (Waitt and Levin 1993, Schlichting 1989, Gianoli 2004, García-Verdugo et al. 2009). However, the stress-induced decanalization theory (Gibson 2009) suggest that new mutations or stressful environments may disrupt fine-tuned connections in a transcriptional network (Lea et al. 2019). Notably, decanalization has been hypothesized to explain complex traits and human disease (Hu et al. 2016). Several studies have shown that environmental perturbation may lead to declining co-expression (Southworth et al. 2009, Anglani et al. 2014).

3.3 Why does the regulatory score perform better than differential expression alone?

We identified a number of cases for which transcription factors fail to exhibit significant differential expression but have a high differential regulatory score (Fig. 4). The different detection thresholds for these two approaches to identify stress-responsive transcription factors might arise due to several molecular processes. First, increased binding frequency might be achieved by altering the abundance of TFs (i.e., differential expression of regulators). Indeed, we found that differential expression can partially explain the differential binding in an additional dataset with ChIP-seq (Fig. S4). With a illustrative model (See Materials and Methods, Fig. 2C and F), the differential expression 'pushes' the abundance of an activator to the active region. Second, the increased binding frequency might arise due to an alteration of binding affinity in response to environmental perturbations, e.g. by the activation of a heat shock TF or its cofactors, or due to the accessibility of corresponding *cis* element in target genes. For instance, the monomeric yeast HSF1 under normal condition has little affinity for the Heat Shock Element (HSE). After conversion into its trimeric and active form at elevated temperature, binding affinity for target DNA motifs is enhanced by several orders of magnitude (Wu 1995). Strictly speaking, an increase in binding affinity without altering the abundance of the activator cannot be detected by co-expression based methods. Thus, we propose that the reason to afford a better performance for the regulatory score in detecting responsive and key regulators is that the expression level of a causal regulator may be not necessarily significantly changed upon the environmental perturbation response — once the abundance of an activator or repressor reaches the threshold of the active region, the dynamic correlation can capture the responsive signal.

3.4 Identifying the molecular basis of stress response

Plant response to many environmental stressors is physiologically complex (Bohnert et al. 1995) and often species-specific (Bouzid et al. 2019). Response to soil drying (drought) is representative of this complex and variable etiology, with considerable among and even within species diversity in the proximate mechanisms of response. Despite the apparently poor conservation of function among orthologs of evolutionarily distant species (Nehrt et al. 2011), newly sequenced genomes are often annotated functionally by gene sequence identity to distantly related reference genomes such as *Arabidopsis thaliana*. While gene expression data generated from controlled experiments on the species of interest can help to improve the accuracy of gene annotations and therefore functional inference, many important functional relationships can be missed by the contrasting environment RNA-Seq studies that are most often employed for annotation (e.g. (Kilian et al. 2007)). The approach presented here provides the means to identify candidate regulatory genes without the experimental challenge of ChIP-Seq, mutant screens, or protein-protein interaction approaches that are commonly employed in model species.

3.5 Large scale dynamic regulatory activity revealed transient response of the transcriptional network

The present work provides a novel data integration pipeline for characterizing context-specific regulatory network and shed light on global dynamic insights into gene regulation. Detecting transient TF-target interaction has long been a challenge in genomics; it is crucial not only to know what gene targets are regulated by a TF but also when (Swift and Coruzzi 2017). Measuring such TF-target dynamics has largely drawn from single molecule approach, which can accurately capture transient interactions within seconds, yet this approach can fail to identify genome-wide targets TFs transiently regulate their targets. While ChIP-seq and ChIP-chip remain the prevailing approaches to learn TF binding genome-wide, whether a TF-DNA binding event detect by ChIP has an effect on transcription and its biological meaning is unknown. Evidence suggests that the overlap of TF-bound and TF-regulated genes (detected by over expression or knock-out mutant lines) is often poor (Swift and Coruzzi 2017). In addition to that, the high cost of genome-wide ChIP-seq has hindered its wide use. Here, instead, our present work suggests that the transient TF-target activity can be characterized by dynamic temporal correlation between a TF and target. As we have shown for heat shock conditions, the differential regulatory activity presents a useful complementary approach for detecting responsive TFs over the traditional differential expression method.

Our results confirmed the previous observations in dynamic regulatory maps as shown in (Luscombe et al. 2004), which used a trace-back algorithm to identify active regulatory links in multiple conditions by integrating gene expression data. By studying the topological feature of active sub-networks under multiple conditions, (Luscombe et al. 2004) showed that environmental perturbations enable fast signal propagation (e.g., with short regulatory cascades and paths). Our results reflect another aspect of fast propagation by showing a faster rate of signal propagation with shorter time lag (Fig. S2).

3.6 Methodology

It is tempting to use ChIP-Chip to characterize the dynamic activity of TRNs in response of environmental cues. However, ChIP-Seq is often locally used in the specific regulon assayed for few known regulators (Song et al. 2016). To make it global and genome-wide, our presented method proposed a multi-omic approach, utilizing a combination of network prior and temporal transcriptome data to infer dynamic activities of TRNs as a proxy to prioritize key regulators. Different structure priors have their unique biases (Greenfield et al. 2013), our analysis validated the feasibility of using static chromatin accessibility data with a TF motif database (Wilkins et al. 2016) to represent potential regulatory links. Although our method relies on a network prior, which may constrain its application in non-model system where the most accessible data is high-throughput mRNA-seq, we believe that a priori transcriptional regulatory network is necessary as the pairwise co-expression network doesn't reflect the causal relationship between two transcripts. With the advent of cost-effective technologies (i.e., [ATAC-seq]) for measuring chromatin accessibility, and the ever-expanding number of reference genomes available for medically, ecologically, and agriculturally relevant species, expression and accessibility data are now becoming available in more time-coursed experiments (Liu et al. 2017, Su et al. 2017, Ramirez et al. 2017, Alvarez-Fernandez et al. 2020, Bechtold et al. 2016).

4 Materials and Methods

4.1 Data retrieval

We utilized a TRN prior previously constructed by Wilkins et al. (2016), which was obtained from the integration of ATAC-seq (Assay for Transposase-Accessible Chromatin using sequencing) data and the CIS-BP database of TF binding motif (Weirauch et al. 2014). We elected not to use their complete "Environmental Gene Regulatory Influence Network" because the estimation of the final network relied on information from mRNA-seq time series data; the analyses presented here represent a unique approach to analyzing these data. Genes that had corresponding **cis**-regulatory elements of TF in a region of open chromatin in their promoter regions are identified as the target gene for a given TF (Wilkins et al. 2016). Note that the open accessible regulatory region derived from ATAC-seq of rice leaves remained stable across multiple environmental conditions in the Wilkins et al study. In total, this "network prior" has 38,137 interactions: 357 TFs were inferred to interact with 3240 target genes. Interactions can be between TFs and non-transcription factor targets, or between two TFs.

The RNA-seq data derive from chamber-grown plants and were retrieved from the Gene Expression Omnibus (`www.ncbi.nlm.nih.gov/geo/`) under accession number GSE74793. This dataset comprises time-coursed libraries for four rice cultivars exposed to control (benign), heat shock, and water deficit conditions. Samples were collected with 15 min intervals for up to 4h for each of the treatments; specifically, 18 time points for controlled conditions; 9 time points for drought treatment; and 16 time points for heat treatment. Here we used a time window of nine time points in each condition for analysis. TF family annotations were downloaded from the Plant TF database `http://plntfdb.bio.uni-potsdam.de/v3.0/downloads.php?sp_id=OSAJ`, from which Heat Shock Factors are those with the TF family label "HSF". "Known" drought-related TFs were obtained from `https://funricegenes.github.io/` in June 2020 using the search terms "drought", "ABA", and "drought tolerance".

4.2 Regulatory score for regulator-target pair

The expression relationship between genes in time series often presents itself by a staggered relationship because of time-lag effect. Traditional correlation coefficients (e.g., Pearson correlation) cannot account for the staggered relationship between a regulator and a target. To better measure the regulatory strength between interacting transcripts, we propose to use max cross-correlation over the time course with a direction constraint (from regulator to target) for

evaluating the regulatory status. Consider two discrete time series denoting $f(t)$ (regulator) and $g(t)$ (target), both of length of N number of time points, the cross-correlation function is defined as:

$$S_{f,g}(\tau) = \frac{1}{N-\tau} \sum_{n=0}^{N-|\tau|-1} \tilde{f}(n+\tau)\tilde{g}(n) \qquad \tau \geq 0$$

$$S_{f,g}(\tau) = S_{g,f}(-\tau) \qquad \tau < 0$$

where the $\tilde{f}(n)$ is a normalized time series (zero mean, unit variance). The regulatory score is defined as the maximum cross correlation $S_{f,g}(\tau)$ under condition of $m \leq \tau \leq 0$, where m is the max delay. The time delay that possesses the max correlation is defined as τ_{reg}, representing the approximate time delay that occurs between a given regulatory-target pair. The max delay is set as 2, which in the current dataset represents a 30 min time interval. We chose a 30 minute interval as the max delay period because several studies have reported that the time duration of transcription and translation is within 30 min (Hirata et al. 2002, Lewis 2003). For comparison between regulatory score distributions under multiple conditions, we used a Kolmogorov-Smirnov test using the *ks.test* function in *R*. Note that it is unknown to what extent of the temporal resolution the present method is effective in revealing the transient dynamic of the regulatory activity (the sampling interval in our dataset is 15 min).

4.3 Simulations for the minimal activation model

To illustrate the potential bias in capturing changing regulatory activities by using population level correlation, we implement simulation through a minimal activation model. The rate of production of TF X and gene Y (Fig. 2A) is described by the following equation:

$$[\dot{X}] = \frac{\beta_x S_x^n}{1+S_x^n} + \beta_{basal} - d_x[X], \tag{1}$$

$$[\dot{Y}] = \frac{\beta_y (\frac{[X]}{k_x})^n}{1+(\frac{[X]}{k_x})^n} + \beta_{basal} - d_y[Y]. \tag{2}$$

$[X]$ and $[Y]$ denote the mRNA concentrations of TF X and gene Y respectively. TF X affects the transcription of gene Y. The regulated expression of genes is represented by Hill function with cooperativity exponent n. Each transcript is assumed to degrade at a rate proportional to its own concentration (d_x and d_y). Assume that the basal synthesis rate for X and Y is constant and equal with β_{basal}. β_y can be taken as the maximum strength of regulations. The stochastic dynamics of the system are implemented through Gillespie stochastic simulation

algorithm (Gillespie 1977). The Hill function has the non-linear and ultrasensitivity property. By tuning the binding affinity K_x and the Hill coefficient n in models of gene regulation, we can manipulate the active range of regulatory interactions. A set of parameters including the induction signal strength S_x are determined to enable two regulatory regimes (Fig. 2C and 2F). Two types of perturbation imposed on cells at steady state are simulated, including press and pulse perturbations (Fig 2B). The press perturbation maintains the external signal at a certain high level throughout the time course, whereas the pulse perturbation indicates a discrete, transient induction of the external signal. We assume the external perturbation modulates the gene expression dynamics by the signal S_x.

Temporal dynamics of TF X and gene Y in a population with 100 cells are simulated. The cross-correlation function is calculated for the bulk time series of X and Y, which are the population average of their expression level over the time course, whereas the population-level Pearson's correlation coefficient (PCC) is calculated at each time point during the simulation.

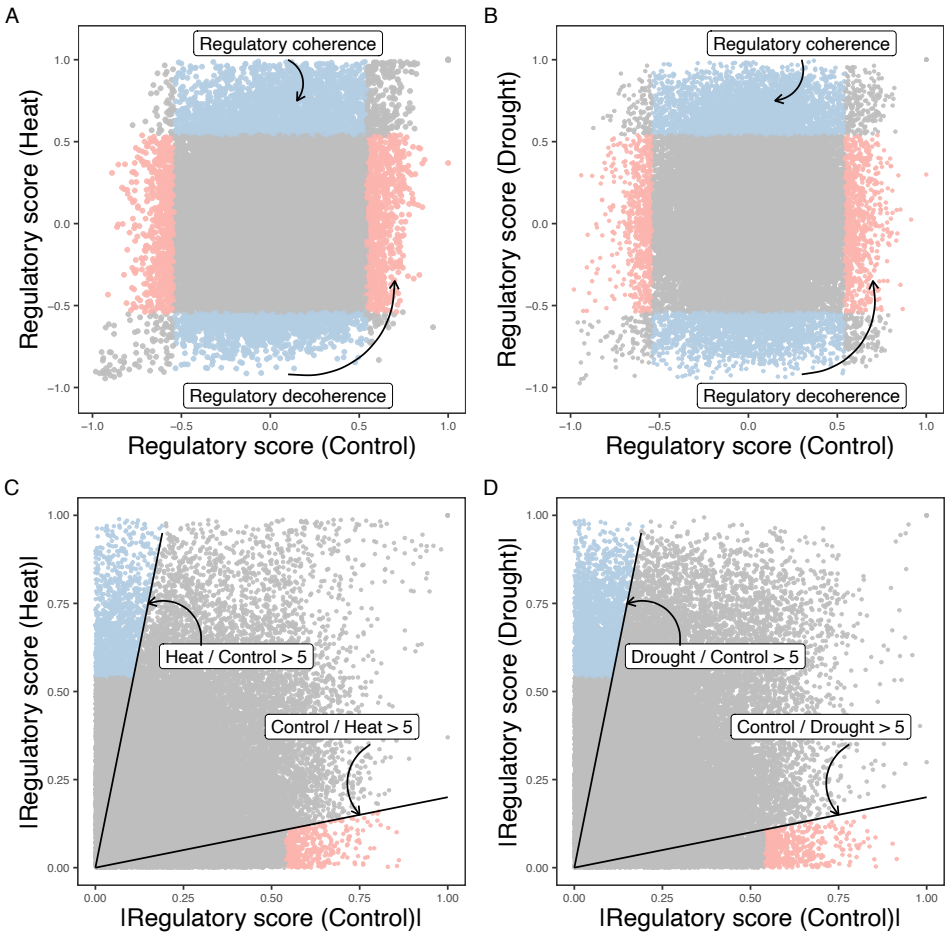

Figure 1: **Environmental perturbations lead to more regulatory coherence. A.** Comparison of the regulatory score for each regulator-target pair under control condition against heat condition. **B.** Comparison of the regulatory score for each regulator-target pair under control condition against soil drying condition. The regulator-target pairs that are not significant on both conditions are in grey, for which the cutoff is 0.54. The red and blue label highlight the pairs that showing regulatory decoherence and regulatory coherence, respectively. **C.** and **D.** show the absolute value of each regulator-target pair with solid lines indicating ratio between regulatory scores under control and perturbed conditions is larger than 5.

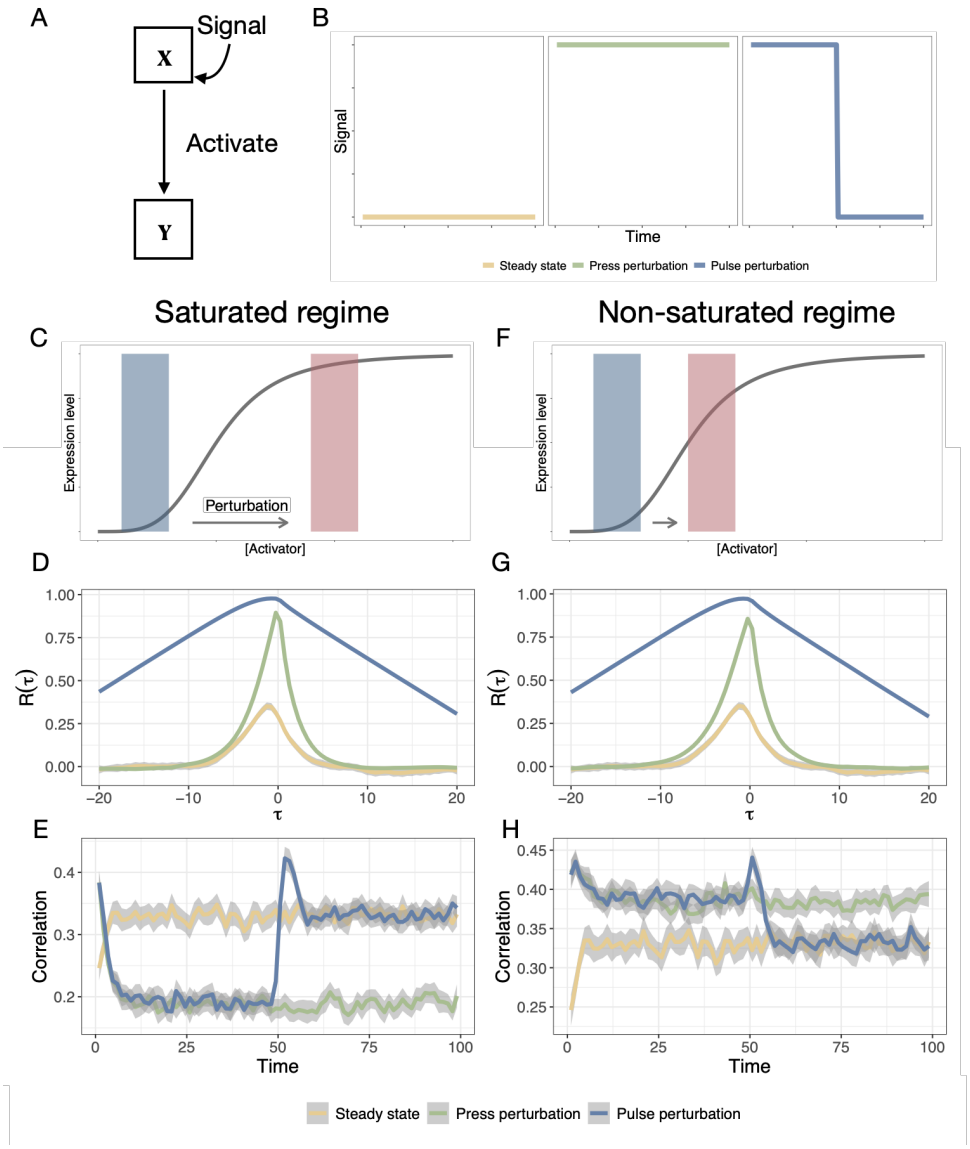

Figure 2: **Illustrated examples from stochastic simulation indicate the robustness of using temporal correlation to characterize the environmentally induced links.** The population level correlation may lead to false negatives in detecting environmentally induced links. **A**. A schematic illustration of minimal activation model and **B**. input signals corresponding to three scenarios. **C - E**. The cross correlation function and population-level correlation between activator X and target Y under a saturated regime. The cross correlation function robustly reveal a peak in response to perturbations while the perturbation may lead to reduction of correlation when using population correlation over the time course. **F - H**. The cross correlation function and population-level correlation between activator X and target Y under a non-saturated regime. Under a non-saturated regime, both the population correlation and temporal correlation can detect elevated level of coexpression. Colors represent three different types of external environmental conditions which lead to internal signaling (Steady state, press perturbation and pulse perturbation). $R(\tau)$ is the cross correlation function with τ indicating the time delay. Note that the perturbation is imposed at $t = 0$ in **E** and **H**.

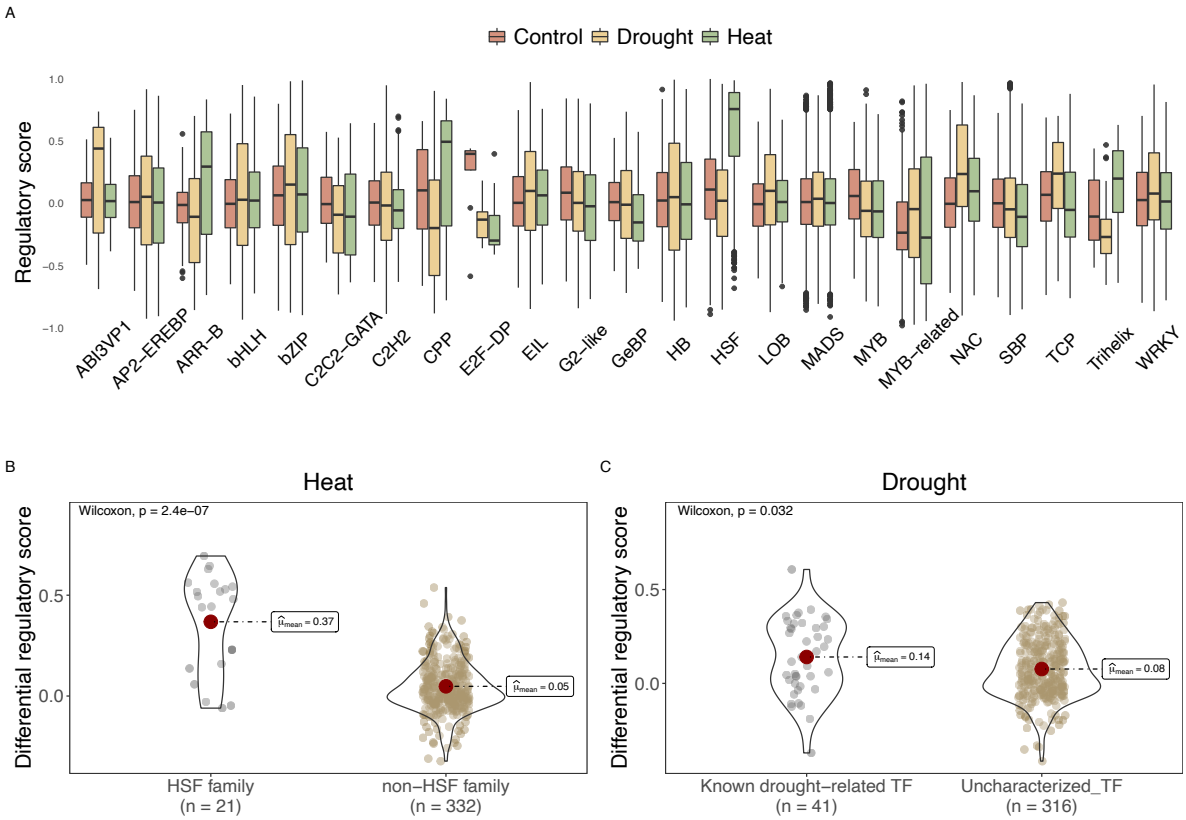

Figure 3: **Responsive regulators drive temporal regulatory coherence. A.** The regulatory score for regulator-target pairs under multiple conditions, grouped by TF families. Each regulator-target pair is assigned a TF family by regulator. **B.** The heat differential regulatory score for regulators in the network prior. **C.** The drought differential regulatory score for regulators in the network prior. Known drought-related TFs were obtained from `https://funricegenes.github.io/`, where genes linked with keywords "drought", "ABA", and "drought tolerance" were extracted. The differential regulatory score is calculated as the averaged score changed across cutoff 0.6 for each regulator.

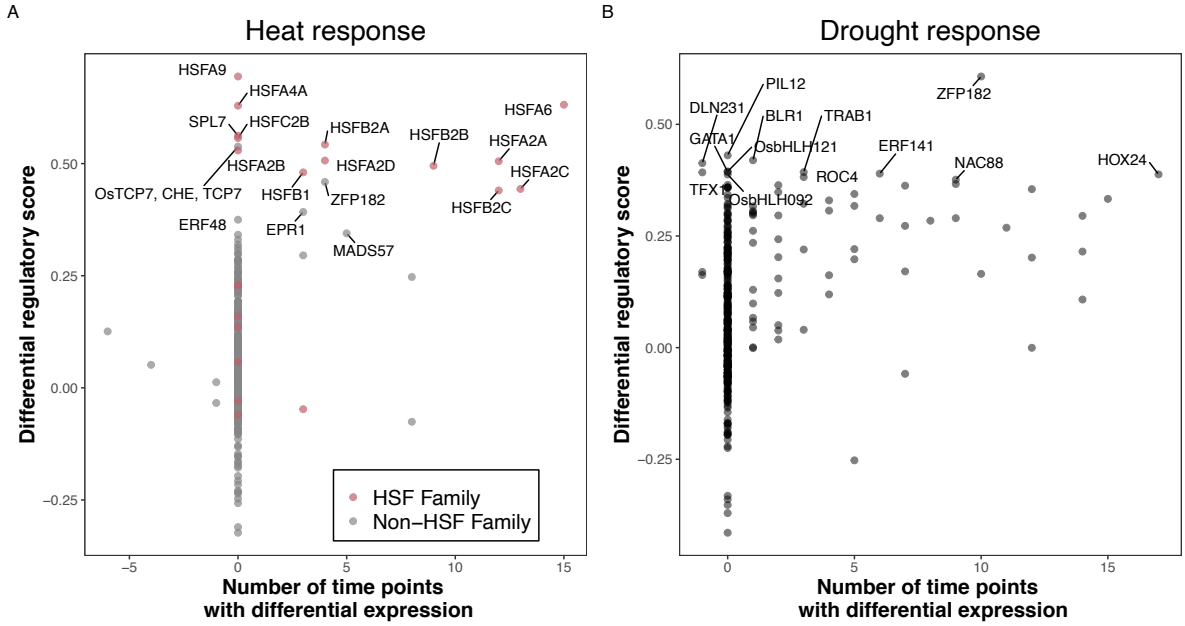

Figure 4: **The regulatory score performs better than the traditional differential expression approach in detecting responsive regulators.** The comparison of heat (A) and drought (B) differential expression strength (the number of time points showing differential expression from the original Wilkins et al. analysis) versus differential regulatory score. Salmon points denote the Heat Shock Family (HSF) regulators.

References

Äijö, T., Butty, V., Chen, Z., Salo, V., Tripathi, S., Burge, C. B., Lahesmaa, R., and Lähdesmäki, H. 2014. Methods for time series analysis of rna-seq data with application to human th17 cell differentiation. Bioinformatics 30:i113–i120.

Alvarez-Fernandez, R., Penfold, C. A., Galvez-Valdivieso, G., Exposito-Rodriguez, M., Stallard, E. J., Bowden, L., Moore, J. D., Mead, A., Davey, P. A., Matthews, J. S., et al. 2020. Time series transcriptomics reveals a bbx32-directed control of dynamic acclimation to high light in mature arabidopsis leaves. bioRxiv .

Amar, D., Safer, H., and Shamir, R. 2013. Dissection of regulatory networks that are altered in disease via differential co-expression. PLoS computational biology 9.

Anglani, R., Creanza, T. M., Liuzzi, V. C., Piepoli, A., Panza, A., Andriulli, A., and Ancona, N. 2014. Loss of connectivity in cancer co-expression networks. PloS one 9.

Azizi, E., Carr, A. J., Plitas, G., Cornish, A. E., Konopacki, C., Prabhakaran, S., Nainys, J., Wu, K., Kiseliovas, V., Setty, M., et al. 2018. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. Cell 174:1293–1308.

Bar-Joseph, Z., Gitter, A., and Simon, I. 2012. Studying and modelling dynamic biological processes using time-series gene expression data. Nature Reviews Genetics 13:552–564.

Barzel, B. and Barabási, A.-L. 2013. Network link prediction by global silencing of indirect correlations. Nature biotechnology 31:720–725.

Bechtold, U., Penfold, C. A., Jenkins, D. J., Legaie, R., Moore, J. D., Lawson, T., Matthews, J. S., Vialet-Chabrand, S. R., Baxter, L., Subramaniam, S., et al. 2016. Time-series transcriptomics reveals that agamous-like22 affects primary metabolism and developmental processes in drought-stressed arabidopsis. The Plant Cell 28:345–366.

Bendall, S. C., Davis, K. L., Amir, E.-a. D., Tadmor, M. D., Simonds, E. F., Chen, T. J., Shenfeld, D. K., Nolan, G. P., and Pe'er, D. 2014. Single-cell trajectory detection uncovers progression and regulatory coordination in human b cell development. Cell 157:714–725.

Bhattacharjee, A., Srivastava, P. L., Nath, O., and Jain, M. 2021. Genome-wide discovery of oshox24-binding sites and regulation of desiccation stress response in rice. Plant Molecular Biology 105:205–214.

Bohnert, H. J., Nelson, D. E., and Jensen, R. G. 1995. Adaptations to environmental stresses. The plant cell 7:1099.

Bouzid, M., He, F., Schmitz, G., Häusler, R., Weber, A., Mettler-Altmann, T., and De Meaux, J. 2019. Arabidopsis species deploy distinct strategies to cope with drought stress. Annals of botany 124:27–40.

Briat, C., Gupta, A., and Khammash, M. 2016. Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. Cell systems 2:15–26.

Buchanan, M., Caldarelli, G., De Los Rios, P., Rao, F., and Vendruscolo, M., 2010. Networks in cell biology. Cambridge University Press.

Campbell, K. and Yau, C. 2015. Bayesian gaussian process latent variable models for pseudotime inference in single-cell rna-seq data. bioRxiv page 026872.

Campbell, K. J., Rocha, S., and Perkins, N. D. 2004. Active repression of antiapoptotic gene expression by rela (p65) nf-κb. Molecular cell 13:853–865.

Campbell, K. R. and Yau, C. 2016. Order under uncertainty: robust differential expression analysis using probabilistic models for pseudotime inference. PLoS computational biology 12:e1005212.

Chahrour, M., Jung, S. Y., Shaw, C., Zhou, X., Wong, S. T., Qin, J., and Zoghbi, H. Y. 2008. Mecp2, a key contributor to neurological disease, activates and represses transcription. Science 320:1224–1229.

Chang, K. N., Zhong, S., Weirauch, M. T., Hon, G., Pelizzola, M., Li, H., Huang, S.-s. C., Schmitz, R. J., Urich, M. A., Kuo, D., et al. 2013. Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in arabidopsis. elife 2:e00675.

Chen, Y., Li, C., Yi, J., Yang, Y., Lei, C., and Gong, M. 2020. Transcriptome response to drought, rehydration and re-dehydration in potato. International journal of molecular sciences 21:159.

Cho, S. B., Kim, J., and Kim, J. H. 2009. Identifying set-wise differential co-expression in gene expression microarray data. BMC bioinformatics 10:109.

Choi, J. K., Yu, U., Yoo, O. J., and Kim, S. 2005. Differential coexpression analysis using microarray data and its application to human cancer. Bioinformatics 21:4348–4355.

Cortijo, S., Bhattarai, M., Locke, J. C., and Ahnert, S. E. 2020. Co-expression networks from gene expression variability between genetically identical seedlings can reveal novel regulatory relationships. Frontiers in plant science 11.

Cuadrado, A., Manda, G., Hassan, A., Alcaraz, M. J., Barbas, C., Daiber, A., Ghezzi, P., León, R., López, M. G., Oliva, B., et al. 2018. Transcription factor nrf2 as a therapeutic target for chronic diseases: a systems medicine approach. Pharmacological reviews 70:348–383.

de la Fuente, A. 2010. From 'differential expression'to 'differential networking'–identification of dysfunctional regulatory networks in diseases. Trends in genetics 26:326–333.

Deng, S.-P., Zhu, L., and Huang, D.-S., 2015. Mining the bladder cancer-associated genes by an integrated strategy for the construction and analysis of differential co-expression networks. *in* BMC genomics, volume 16, page S4. BioMed Central.

Dunlop, M. J., Cox, R. S., Levine, J. H., Murray, R. M., and Elowitz, M. B. 2008. Regulatory activity revealed by dynamic correlations in gene expression noise. Nature genetics 40:1493–1498.

Duren, Z., Chen, X., Xin, J., Wang, Y., and Wong, W. H. 2020. Time course regulatory analysis based on paired expression and chromatin accessibility data. Genome research 30:622–634.

Ernst, J., Vainas, O., Harbison, C. T., Simon, I., and Bar-Joseph, Z. 2007. Reconstructing dynamic regulatory maps. Molecular systems biology 3:74.

Feizi, S., Marbach, D., Médard, M., and Kellis, M. 2013. Network deconvolution as a general method to distinguish direct dependencies in networks. Nature biotechnology 31:726–733.

Fukushima, A. 2013. Diffcorr: an r package to analyze and visualize differential correlations in biological networks. Gene 518:209–214.

Fukushima, A., Nishizawa, T., Hayakumo, M., Hikosaka, S., Saito, K., Goto, E., and Kusano, M. 2012. Exploring tomato gene functions based on coexpression modules using graph clustering and differential coexpression approaches. Plant physiology 158:1487–1502.

García-Verdugo, C., Granado-Yela, C., Manrique, E., Rubio de Casas, R., and Balaguer, L. 2009. Phenotypic plasticity and integration across the canopy of olea europaea subsp. guanchica (oleaceae) in populations with different wind exposures. American Journal of Botany 96:1454–1461.

Gargouri, M., Park, J.-J., Holguin, F. O., Kim, M.-J., Wang, H., Deshpande, R. R., Shachar-Hill, Y., Hicks, L. M., and Gang, D. R. 2015. Identification of regulatory network hubs that control lipid metabolism in chlamydomonas reinhardtii. Journal of experimental botany 66:4551–4566.

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D., and Brown, P. O. 2000. Genomic expression programs in the response of yeast cells to environmental changes. Molecular biology of the cell 11:4241–4257.

Gianoli, E. 2004. Plasticity of traits and correlations in two populations of convolvulus arvensis (convolvulaceae) differing in environmental heterogeneity. International Journal of Plant Sciences 165:825–832.

Gibson, G. 2008. The environmental contribution to gene expression profiles. Nature reviews genetics 9:575–581.

Gibson, G. 2009. Decanalization and the origin of complex disease. Nature Reviews Genetics 10:134–140.

Gillespie, D. T. 1977. Exact stochastic simulation of coupled chemical reactions. The journal of physical chemistry 81:2340–2361.

Greenfield, A., Hafemeister, C., and Bonneau, R. 2013. Robust data-driven incorporation of prior knowledge into the inference of dynamic regulatory networks. Bioinformatics 29:1060–1067.

Greenham, K., Guadagno, C. R., Gehan, M. A., Mockler, T. C., Weinig, C., Ewers, B. E., and McClung, C. R. 2017. Temporal network analysis identifies early physiological and transcriptomic indicators of mild drought in brassica rapa. Elife 6:e29655.

Hahn, J.-S., Hu, Z., Thiele, D. J., and Iyer, V. R. 2004. Genome-wide analysis of the biology of stress responses through heat shock transcription factor. Molecular and cellular biology 24:5249–5256.

Harbison, C. T., Gordon, D. B., Lee, T. I., Rinaldi, N. J., Macisaac, K. D., Danford, T. W., Hannett, N. M., Tagne, J.-B., Reynolds, D. B., Yoo, J., et al. 2004. Transcriptional regulatory code of a eukaryotic genome. Nature 431:99–104.

Hirata, H., Yoshiura, S., Ohtsuka, T., Bessho, Y., Harada, T., Yoshikawa, K., and Kageyama, R. 2002. Oscillatory expression of the bhlh factor hes1 regulated by a negative feedback loop. Science 298:840–843.

Hu, J. X., Thomas, C. E., and Brunak, S. 2016. Network biology concepts in complex disease comorbidities. Nature Reviews Genetics 17:615.

Huang, J., Sun, S., Xu, D., Lan, H., Sun, H., Wang, Z., Bao, Y., Wang, J., Tang, H., and Zhang, H. 2012. A tfiiia-type zinc finger protein confers multiple abiotic stress tolerances in transgenic rice (oryza sativa l.). Plant molecular biology 80:337–350.

Joshi, R., Wani, S. H., Singh, B., Bohra, A., Dar, Z. A., Lone, A. A., Pareek, A., and Singla-Pareek, S. L. 2016. Transcription factors and plants response to drought stress: current understanding and future directions. Frontiers in Plant Science 7:1029.

Kafri, R., Levy, J., Ginzberg, M. B., Oh, S., Lahav, G., and Kirschner, M. W. 2013. Dynamics extracted from fixed cells reveal feedback linking cell growth to cell cycle. Nature 494:480–483.

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

Kostka, D. and Spang, R. 2004. Finding disease specific alterations in the co-expression of genes. Bioinformatics 20:i194–i199.

Lan Thi Hoang, X., Du Nhi, N. H., Binh Anh Thu, N., Phuong Thao, N., and Phan Tran, L.-S. 2017. Transcription factors and their roles in signal transduction in plants under abiotic stresses. Current genomics 18:483–497.

Lea, A., Subramaniam, M., Ko, A., Lehtimäki, T., Raitoharju, E., Kähönen, M., Seppälä, I., Mononen, N., Raitakari, O. T., Ala-Korpela, M., et al. 2019. Genetic and environmental perturbations lead to regulatory decoherence. eLife 8:e40538.

Lee, T. I. and Young, R. A. 2013. Transcriptional regulation and its misregulation in disease. Cell 152:1237–1251.

Lewis, J. 2003. Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. Current Biology 13:1398–1408.

Li, P., Yang, H., Wang, L., Liu, H., Huo, H., Zhang, C., Liu, A., Zhu, A., Hu, J., Lin, Y., et al. 2019. Physiological and transcriptome analyses reveal short-term responses and formation of memory under drought stress in rice. Frontiers in genetics 10:55.

Li, W. V. and Li, Y. 2020. sclink: Inferring sparse gene co-expression networks from single-cell expression data. bioRxiv .

Liu, Q., Jiang, C., Xu, J., Zhao, M.-T., Van Bortle, K., Cheng, X., Wang, G., Chang, H. Y., Wu, J. C., and Snyder, M. P. 2017. Genome-wide temporal profiling of transcriptome and open chromatin of early cardiomyocyte differentiation derived from hipscs and hescs. Circulation research 121:376–391.

Luscombe, N. M., Babu, M. M., Yu, H., Snyder, M., Teichmann, S. A., and Gerstein, M. 2004. Genomic analysis of regulatory network dynamics reveals large topological changes. Nature 431:308–312.

Ma, W., Trusina, A., El-Samad, H., Lim, W. A., and Tang, C. 2009. Defining network topologies that can achieve biochemical adaptation. Cell 138:760–773.

Manna, M., Thakur, T., Chirom, O., Mandlik, R., Deshmukh, R., and Salvi, P. 2020. Transcription factors as key molecular target to strengthen the drought stress tolerance in plants. Physiologia Plantarum .

Martín-Trillo, M. and Cubas, P. 2010. Tcp genes: a family snapshot ten years later. Trends in plant science 15:31–39.

Matsumoto, H., Kiryu, H., Furusawa, C., Ko, M. S., Ko, S. B., Gouda, N., Hayashi, T., and Nikaido, I. 2017. Scode: an efficient regulatory network inference algorithm from single-cell rna-seq during differentiation. Bioinformatics 33:2314–2321.

Moignard, V., Woodhouse, S., Haghverdi, L., Lilly, A. J., Tanaka, Y., Wilkinson, A. C., Buettner, F., Macaulay, I. C., Jawaid, W., Diamanti, E., et al. 2015. Decoding the regulatory network of early blood development from single-cell gene expression measurements. Nature biotechnology 33:269–276.

Müller, M., Seifert, S., Lübbe, T., Leuschner, C., and Finkeldey, R. 2017. De novo transcriptome assembly and analysis of differential gene expression in response to drought in european beech. PloS one 12:e0184167.

Munsky, B., Neuert, G., and Van Oudenaarden, A. 2012. Using gene expression noise to understand gene regulation. Science 336:183–187.

Nakamura, Y., Kato, T., Yamashino, T., Murakami, M., and Mizuno, T. 2007. Characterization of a set of phytochrome-interacting factor-like bhlh proteins in oryza sativa. Bioscience, Biotechnology, and Biochemistry 71:1183–1191.

Neef, D. W., Jaeger, A. M., and Thiele, D. J. 2011. Heat shock transcription factor 1 as a therapeutic target in neurodegenerative diseases. Nature reviews Drug discovery 10:930–944.

Neef, D. W., Turski, M. L., and Thiele, D. J. 2010. Modulation of heat shock transcription factor 1 as a therapeutic target for small molecule intervention in neurodegenerative disease. PLoS biology 8.

Nehrt, N. L., Clark, W. T., Radivojac, P., and Hahn, M. W. 2011. Testing the ortholog conjecture with comparative functional genomic data from mammals. PLoS Comput Biol 7:e1002073.

Ohama, N., Sato, H., Shinozaki, K., and Yamaguchi-Shinozaki, K. 2017. Transcriptional regulatory network of plant heat stress response. Trends in plant science 22:53–65.

Parsana, P., Ruberman, C., Jaffe, A. E., Schatz, M. C., Battle, A., and Leek, J. T. 2019. Addressing confounding artifacts in reconstruction of gene co-expression networks. Genome biology 20:1–6.

Piggot, P. J. and Hilbert, D. W. 2004. Sporulation of bacillus subtilis. Current opinion in microbiology 7:579–586.

Ramirez, R. N., El-Ali, N. C., Mager, M. A., Wyman, D., Conesa, A., and Mortazavi, A. 2017. Dynamic gene regulatory networks of human myeloid differentiation. Cell systems 4:416–429.

Rastogi, A., Maheswari, U., Dorrell, R. G., Vieira, F. R. J., Maumus, F., Kustka, A., McCarthy, J., Allen, A. E., Kersey, P., Bowler, C., et al. 2018. Integrative analysis of large scale transcriptome data draws a comprehensive landscape of phaeodactylum tricornutum genome and evolutionary origin of diatoms. Scientific reports 8:1–14.

Reid, J. E. and Wernisch, L. 2016. Pseudotime estimation: deconfounding single cell time series. Bioinformatics 32:2973–2980.

Schlichting, C. D. 1989. Phenotypic integration and environmental change. BioScience 39:460.

Smith, H. 1990. Signal perception, differential expression within multigene families and the molecular basis of phenotypic plasticity. Plant, Cell & Environment 13:585–594.

Song, L., Huang, S.-s. C., Wise, A., Castanon, R., Nery, J. R., Chen, H., Watanabe, M., Thomas, J., Bar-Joseph, Z., and Ecker, J. R. 2016. A transcription factor hierarchy defines an environmental stress response network. Science 354:aag1550.

Southworth, L. K., Owen, A. B., and Kim, S. K. 2009. Aging mice show a decreasing correlation of gene expression within genetic modules. PLoS genetics 5.

Spampanato, C., Feeney, E., Li, L., Cardone, M., Lim, J.-A., Annunziata, F., Zare, H., Polishchuk, R., Puertollano, R., Parenti, G., et al. 2013. Transcription factor eb (tfeb) is a new therapeutic target for pompe disease. EMBO molecular medicine 5:691–706.

Specht, A. T. and Li, J. 2017. Leap: constructing gene co-expression networks for single-cell rna-sequencing data using pseudotime ordering. Bioinformatics 33:764–766.

Su, Y., Shin, J., Zhong, C., Wang, S., Roychowdhury, P., Lim, J., Kim, D., Ming, G.-l., and Song, H. 2017. Neuronal activity modifies the chromatin accessibility landscape in the adult brain. Nature neuroscience 20:476.

Sun, X., Zhang, J., and Nie, Q. 2021. Inferring latent temporal progression and regulatory networks from cross-sectional transcriptomic data of cancer samples. PLOS Computational Biology 17:e1008379.

Swift, J. and Coruzzi, G. M. 2017. A matter of time—how transient transcription factor interactions create dynamic gene regulatory networks. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms 1860:75–83.

Todaka, D., Nakashima, K., Maruyama, K., Kidokoro, S., Osakabe, Y., Ito, Y., Matsukura, S., Fujita, Y., Yoshiwara, K., Ohme-Takagi, M., et al. 2012. Rice phytochrome-interacting factor-like protein ospil1 functions as a key regulator of internode elongation and induces a morphological response to drought stress. Proceedings of the National Academy of Sciences 109:15947–15952.

Toledo, F. and Wahl, G. M. 2006. Regulating the p53 pathway: in vitro hypotheses, in vivo veritas. Nature Reviews Cancer 6:909–923.

Van Kleunen, M. and Fischer, M. 2005. Constraints on the evolution of adaptive phenotypic plasticity in plants. New Phytologist 166:49–60.

Varala, K., Marshall-Colón, A., Cirrone, J., Brooks, M. D., Pasquino, A. V., Léran, S., Mittal, S., Rock, T. M., Edwards, M. B., Kim, G. J., et al. 2018. Temporal transcriptional logic of dynamic regulatory networks underlying nitrogen signaling and use in plants. Proceedings of the National Academy of Sciences 115:6494–6499.

von Koskull-Döring, P., Scharf, K.-D., and Nover, L. 2007. The diversity of plant heat stress transcription factors. Trends in plant science 12:452–457.

Waitt, D. E. and Levin, D. A. 1993. Phenotypic integration and plastic correlations in phlox drummondii (polemoniaceae). American Journal of Botany 80:1224–1233.

Wang, H., Wang, H., Shao, H., and Tang, X. 2016. Recent advances in utilizing transcription factors to improve plant abiotic stress tolerance by transgenic technology. Frontiers in plant science 7:67.

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. 2004. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. Trends in plant science 9:244–252.

Weirauch, M. T., Yang, A., Albu, M., Cote, A. G., Montenegro-Montero, A., Drewe, P., Najafabadi, H. S., Lambert, S. A., Mann, I., Cook, K., et al. 2014. Determination and inference of eukaryotic transcription factor sequence specificity. Cell 158:1431–1443.

Wilkins, O., Hafemeister, C., Plessis, A., Holloway-Phillips, M.-M., Pham, G. M., Nicotra, A. B., Gregorio, G. B., Jagadish, S. K., Septiningsih, E. M., Bonneau, R., et al. 2016. Egrins (environmental gene regulatory influence networks) in rice that function in the response to water deficit, high temperature, and agricultural environments. The Plant Cell 28:2365–2384.

Windram, O., Madhou, P., McHattie, S., Hill, C., Hickman, R., Cooke, E., Jenkins, D. J., Penfold, C. A., Baxter, L., Breeze, E., et al. 2012. Arabidopsis defense against botrytis cinerea: chronology and regulation deciphered by high-resolution temporal transcriptomic analysis. The Plant Cell 24:3530–3557.

Wray, G. A., Hahn, M. W., Abouheif, E., Balhoff, J. P., Pizer, M., Rockman, M. V., and Romano, L. A. 2003. The evolution of transcriptional regulation in eukaryotes. Molecular biology and evolution 20:1377–1419.

Wu, C. 1995. Heat shock transcription factors: structure and regulation. Annual review of cell and developmental biology 11:441–469.

Yan, Q., Wu, F., Yan, Z., Li, J., Ma, T., Zhang, Y., Zhao, Y., Wang, Y., and Zhang, J. 2019. Differential co-expression networks of long non-coding rnas and mrnas in cleistogenes songorica under water stress and during recovery. BMC plant biology 19:1–19.

Yeung, E., van Veen, H., Vashisht, D., Paiva, A. L. S., Hummel, M., Rankenberg, T., Steffens, B., Steffen-Heins, A., Sauter, M., de Vries, M., et al. 2018. A stress recovery signaling network for enhanced flooding tolerance in arabidopsis thaliana. Proceedings of the National Academy of Sciences 115:E6085–E6094.

Yuan, W., Beaulieu-Jones, B. K., Yu, K.-H., Lipnick, S. L., Palmer, N., Loscalzo, J., Cai, T., and Kohane, I. S. 2021. Temporal bias in case-control design: preventing reliable predictions of the future. Nature Communications 12:1107. URL https://doi.org/10.1038/s41467-021-21390-2.

Zander, M., Lewsey, M. G., Clark, N. M., Yin, L., Bartlett, A., Guzmán, J. P. S., Hann, E., Langford, A. E., Jow, B., Wise, A., et al. 2020. Integrated multi-omics framework of the plant response to jasmonic acid. Nature Plants 6:290–302.