

Article

Network Biology Analyses and Dynamic Modeling of Gene Regulatory Networks Under Drought Stress Reveal Major Transcriptional Regulators in Arabidopsis

Nilesh Kumar ^{1,†}, Bharat K. Mishra ^{1,†}, Jinbao Liu ^{1,†}, Binoop Mohan ¹, Karolina M. Pajerowska-Mukhtar ^{1,‡} and M. Shahid Mukhtar ^{1,2,3,‡,*}

¹Department of Biology, 464 Campbell Hall, 1300 University Boulevard, University of Alabama at Birmingham, Alabama 35294, USA

²Nutrition Obesity Research Center, 1675 University Boulevard, University of Alabama at Birmingham, Birmingham, AL 35294, USA

³Department of Surgery, 1808 7th Ave S, University of Alabama at Birmingham, Birmingham, AL 35294, USA

[†]These authors contributed equally to this work and share first authorship

[‡]These authors contributed equally to this work and share senior authorship

* Correspondence: smukhtar@uab.edu;

Abstract: Drought is one of the most serious abiotic stressors in the environment, restricting agricultural production by reducing plant growth, development, and productivity. To investigate such a complex and multifaceted stressor and its effects on the plants, a systems biology-based approach is necessitated, entailing the generation of co-expression networks, identification of highly-priority transcription factors (TFs) dynamic mathematical modeling, and computational simulations. Here, we studied a high-resolution drought transcriptome of Arabidopsis. We revealed distinct temporal transcriptional signatures and demonstrated the involvement of specific biological pathways. Generation of a large-scale co-expression network followed by network centrality analyses identified 117 TFs that possess critical properties of hubs, bottlenecks, and high clustering coefficients nodes. Dynamic transcriptional regulatory modeling on integrated TF-target and transcriptome datasets uncovered major transcriptional events during the course of drought stress. Mathematical transcriptional simulations allowed us to ascertain the activation status of major TFs as well as the transcriptional intensity and amplitude of their target genes. Finally, we validated our predictions by providing experimental evidence of gene expression under drought stress for a set of four TFs and their major target genes using qRT-PCR. Taken together, we provided a systems-level perspective on the dynamic transcriptional regulation during drought stress in Arabidopsis and uncovered numerous novel TFs that can potentially be used in future genetic crop engineering programs.

Keywords: co-expression network; water deprivation; drought; systems biology; network centrality; computational simulation; Arabidopsis; transcriptional regulation

1. Introduction

Since the last century, ecological and agricultural systems have been negatively impacted by a wide spectrum of environmental cues including biotic and abiotic stresses. Among all the abiotic stimuli that limit crop yield, drought is the most threatening environmental constraint that adversely affects plant growth and productivity [1–4]. As such, it constitutes a serious threat to animals and crops in practically every corner of the globe by affecting up to 45 percent of the world's agricultural area, which is home to 38 percent of the world's human population [2]. In the United States, drought is increasingly common, costing an average of \$9.6 billion each year, and is thus deemed as one of the most critical economic problems and natural threats [2]. Furthermore, this crisis of crop productivity is further exacerbated by global climate change and future water shortages [3]. Since water is essential at all stages of growth and development, a subtle variation in water

potential during critical growth phases can adversely impact the yield of agronomically important crops [2,5,6]. For instance, drought-related water scarcity can reduce 40% and 50% of wheat and tomato production, respectively [5,7]. Likewise, it was projected that climate change-related water deficiency can reduce potato production by 18-32% in the next four decades [4]. Besides directly impacting crop yield, drought can induce nutrient deficiencies owing to a slower rate of mineral diffusion from the soil matrix to the roots [8]. Furthermore, drought stress elicits a wide range of physiological, biochemical, and metabolic injuries in plants leading to cellular damage and reduction in crop production and quality [2,9]. Therefore, it is imperative to understand the underlying molecular mechanisms of plant drought stress tolerance.

While diverse plant organs and tissues including root, shoot, and leaves respond to drought stress in both a generalized and specialized manner [10], roots constitute the primary sensors detecting drought-related stresses. At the onset of drought stress, the modification of root architecture, as well as regulation of root to shoot water traffic, are initiated to maximize the water uptake [8,10,11]. To avoid non-stomatal water loss in the aerial parts of plants, drought induces structural and compositional cuticular modifications that augment water barrier property [11]. Furthermore, stomata located in the epidermis of the leaves are closed to minimize excessive water loss through transpiration. Such physiological changes in the whole plant are coordinated with the alteration in the cellular concentration of diverse phytohormones, particularly the production of abscisic acid (ABA) and ethylene (ET) [10,12]. This leads to the activation of intricate hormonal signaling networks that are controlled by functional and regulatory proteins and manifested by global transcriptomic reprogramming [13–15]. Towards this, several major plant transcription factor (TF) families have been characterized in response to drought. This includes AREB/ABF (ABA-responsive element-binding protein/ABA-binding factor), DREB (dehydration-responsive element-binding), AP2/ERF (APETALA2/ethylene-responsive element-binding factor), bZIP (basic leucine zipper), NAC (NAM, ATAF1/2, CUC), MYB (myeloblastosis oncogene), and bHLH (basic helix-loop-helix proteins) [10,12–17]. While a wealth of drought-related transcriptomic information is available for the model plant *Arabidopsis*, it remains to be determined how temporal transcriptomic differences are governed by master regulators including major TFs. To establish a comprehensive understanding of complex and dynamic host transcriptional regulatory responses to drought stress, a systems-level approach is necessitated that encompasses systematic, time-scale transcriptional response analyses in *Arabidopsis* that could be later applied to agronomically important crop plants.

A Systems Biology-based investigation of transcriptional regulatory networks (TRNs) entails the generation of genome-wide expression networks, identification of highly influential nodes, as well as dynamic mathematical modeling, computational simulations, and probabilistic assessments of transcription [18–21]. Regulatory networks comprise transcriptional components, referred to as “nodes”, and regulatory interactions between nodes, termed “edges” [20,22]. The characteristics of such transcriptional relationship constitute either direct TF-target interaction or co-expression of highly correlated genes [10,13,14,18]. Network biology, a branch of systems biology, is critical in deciphering biological information of individual nodes or edges and the regulatory networks as a whole [20,23–25]. Henceforth, topological features of a system and network architectural analyses can be employed to identify novel molecular players and regulators in various regulatory networks in diverse systems [14,26–29]. Among these network features, modules (clusters of highly correlated genes) and hubs (high degree; the number of edges of a node) have been used to predict major regulators in *Arabidopsis* and tomato [14,30]. Modeling dynamic transcriptional processes in TRNs is challenging, but it can hold a great promise to comprehend how plants respond to water scarcity. For instance, in a small-scale genetics-based study, an *Arabidopsis* TF ANAC016 was shown to coordinate with AREB1 and NAP (NAC-like activated by AP3/PI) TFs through a trifurcate feed-forward regulatory circuitry in response to drought [31]. iDREM (interactive Dynamic Regulatory

Events Miner) reconstructs dynamic regulatory networks at global levels that can potentially provide information on the function and timing of master regulators [32]. It utilizes time-course RNA-seq and static TF-target datasets and highlights major bifurcation events where the expression of a subset of genes diverges from the rest of the genes and identifies master regulators potentially responsible for the transcriptional events [32,33]. SQUAD (Standardized Qualitative Dynamical systems) performs steady-state studies and dynamic mathematical simulations of TFs and their interacting targets [34]. It combines Boolean and ordinary differential equation (ODE) models to execute such mathematical simulations [34]. Taken together, a systems biology-based approach allows the construction of global drought-responsive TRNs and identifies master regulators that play critical roles throughout different stages of drought.

Here, we identified time point-specific unique and shared Arabidopsis transcriptome signatures in response to drought. Moreover, we highlighted the biological pathways pertinent to these temporal transcriptomes. We constructed a global co-expression network in Arabidopsis to determine highly significant modules that are implicated in drought. Our network architectural and centrality analyses identified high-priority genes and 117 TFs. Integration of TF-target and time-course RNA-seq dataset allowed us to identify 25 major TFs that are implicated in “water deprivation response” throughout the time of drought stress. We selected six of these novel TFs, which possess either of three centrality features i.e., hubs, bottleneck as well as high clustering coefficient (CC), and performed mathematical simulations that allow us to decipher the amplitude and duration of target gene expression exerted by these six TFs. To support these computational transcriptional predictions, we provided experimental evidence of the TF-target relationship in a time-course drought qRT-PCR experiment using Arabidopsis plants. Overall, we provided a comprehensive dynamic model of the transcriptional regulatory network in Arabidopsis during drought stress and discovered several novel TFs that can be implicated in future genetic engineering crop enhancement programs.

2. Results

2.1. The temporal transcriptional landscape of Arabidopsis during drought stress identifies unique transcriptome signatures and specific biological pathways.

To reveal the drought-related transcriptional landscape in Arabidopsis, we built a systems biology-based pipeline (Figure S1). Towards this, we first analyzed a high-resolution mRNA expression dataset (GSE76827) that pertains to drought stress treatment at day 0, 1, 3, 5, 7 and 9 (Figure 1A, Table S1). We observed only a few hundred differentially expressed genes (DEGs) in the early phase of drought stress i.e., 42 and 947 DEGs for day 1 and day 3, respectively (Figure 1B, Table S1). However, we discovered a drastic change in gene expression as the drought stress progressed to day 5 (3,924 DEGs) and day 7 (9,285 DEGs) with the highest numbers of DEGs at day 9 (13,614 DEGs) (Figure 1B, Table S1). Intriguingly, we found that similar numbers of genes were up- or down-regulated at all data points. For instance, 6,702 and 6,912 genes were up- and down-regulated on day 9 of drought stress, respectively (Figure 1B, Table S1). To uncover both the shared and distinct temporal transcriptional landscapes for each time point, we performed expression gradient analysis and identified time point-specific unique DEGs. This method of gene expression analysis allowed us to determine numerous novel and unique transcriptional signatures over the course of drought progression. Specifically, we discovered 1, 46, 265, 940, and 5,471 DEGs on day 1, day 3, day 5, day 7, and day 9, respectively (Figure 1C, Table S1). These results indicate that the number of unique genes also gradually increased akin to the cumulative number of DEGs on a given day of drought stress. Compared to day 7, which exhibited only 940 unique transcriptional signatures out of 9,285 DEGs, day 9 of drought stress is manifested with a majority of unique DEGs, i.e. 5,471 out of 13,614 DEGs (Figure 1C, Table S1). Overall, these data suggest that while the amplitude and duration of transcriptional signatures might be different between day 1 and day 7 of drought stress,

the transcriptional landscape is overall very similar for this phase of drought. Moreover, a significant transcriptional switch likely ensued on day 9 (Figure 1C, Table S1). To further understand the biological characteristics of such transcriptional switch and gain insights into the potential pathways that are driving these time point-specific DEGs under drought stress, we performed gene annotation and pathway analysis using Metascape [35]. We found the “response to stimulus” gene ontology (GO) category was enriched on day 3, while day 5 was enriched with “pyrimidine nucleoside metabolic process”, “regulation of cell communication”, “sulfur and amino acid metabolic” and “ion transport” GO categories (Figure S2). Intriguingly, day 7 GO categories were pertinent to cell wall biosynthesis, DNA damage, and cellular stress responses (Figure 1D). Finally, we observed that drought response on day 9 featured GO categories related to catabolic processes such as “cell death” and “autophagy”, “proteolysis”, “vesicle-mediated transport”, “RNA splicing” and “chromatin organization” (Figure 1E). Taken together, our analyses constructed a global, dynamic transcriptional landscape of drought that features major transcriptional events, unique gene signatures, and novel pathways implicated in the Arabidopsis drought responses.

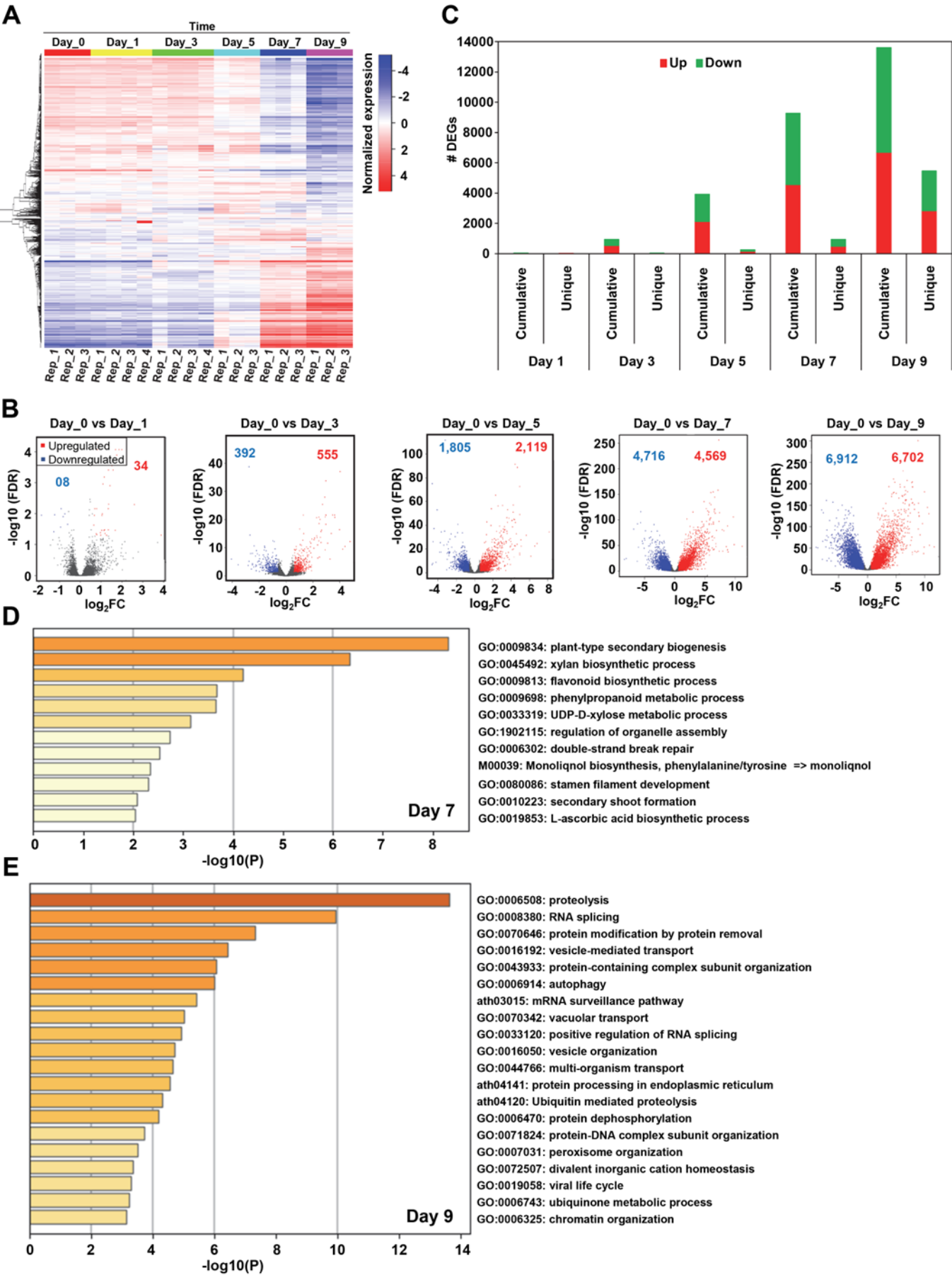


Figure 1. Arabidopsis transcriptome dataset (GSE76827) displays more change in the expression patterns at later sampling days of drought stress. (A) Heatmap of the gene expression profile of Arabidopsis under drought conditions on different sampling days (as indicated). The color bars display the normalized gene expression (Red ≥ 1 , Blue ≥ -1). (B) The total number of uniquely expressed genes on day 1, 3, 5, 7 and 9 are shown in the plots. (C) Differential gene expression is displayed during the drought treatment time frame. The total number of genes significantly differentially expressed were 42, 947, 3924, 9285 and 13,614, respectively, on day 1, 3, 5, 7 and 9 (false discovery rate (FDR) ≤ 0.05 , log2 Fold Change (log2FC) $\geq |0.58|$). Red dots correspond to up-regulated genes, while blue points indicate down-regulated genes. (D) and (E) The gene ontology (GO) enrichment analysis of Arabidopsis unique differentially expressed genes (DEGs) on day 7; D, and day 9; E) after drought stress is presented (p -value ≤ 0.05).

2.2. Network topology analyses of drought-related co-expression network reveal 117 major TFs

A gene co-expression network is a systems biology-based analysis that isolates and highlights various correlation patterns among a set of genes and identifies diverse modules encompassing clusters of highly correlated genes that are potentially involved in related biological pathways. Such analysis can allow the prediction of gene functions as well as the discovery of functional modules, which may provide further insights into the transcriptional gene regulatory mechanisms. Towards generating an Arabidopsis Drought-specific Gene Co-expression Network (ADGCN), we used WGCNA (Weighted gene co-expression network analysis) [36]. Our ADGCN encompasses 9,370 nodes (genes) connected by 402,598 weighted edges (co-expression relationship) (Figure 2A). To understand the nature of interactions within the network and ascertain highly connected modules in ADGCN, we first calculated the degree of distribution and performed a connectivity analysis. We demonstrated that ADGCN exhibits properties of a scale-free network ($R^2 > 0.9$), a power-law degree distribution possessing fewer highly connected nodes with other nodes in a network (Figure S3A). Subsequently, we demonstrated the network connectivity that was indicated by the location of genes in a dendrogram (Figure S3B). Topological Overlap Mapping Metric (TOM) led to the identification of 15 diverse modules (Figure S3B, Table S1). Moving forward, we calculated the network features of ADGCN. We demonstrated that the average clustering coefficient and connectivity of ADGCN were 0.43 and 144, respectively (Figures S2C, D). Subsequently, to retain the most contributing modules in ADGCN, we focused our analyses on five modules that displayed an increased clustering coefficient over the entire network. Furthermore, it's important to note that two of these five modules also possessed heightened connectivity over the entire network (Figures S2C, D). Subsequently, we focused our analyses on five modules that displayed an increased clustering coefficient. It's important to note that two of these five modules also possessed heightened connectivity over the entire network (Figures S2C, D). To provide functional insights into these modules, we next performed pathway analyses (Figure 2A). The "blue" module was enriched with "amino acid homeostasis", "autophagy", "chemotaxis, and "mRNA splicing". Interestingly, these GO categories were the hallmark of unique DEGs of day 9 (Figure 2A). Likewise, we found that the significantly enriched GO terms of the "turquoise" module were "photosynthesis, "biosynthetic process", "mitotic cell cycle", "DNA replication", and "amino acid metabolism", which were associated with the unique DEGs of day 5 and day 7 (Figure 2A).

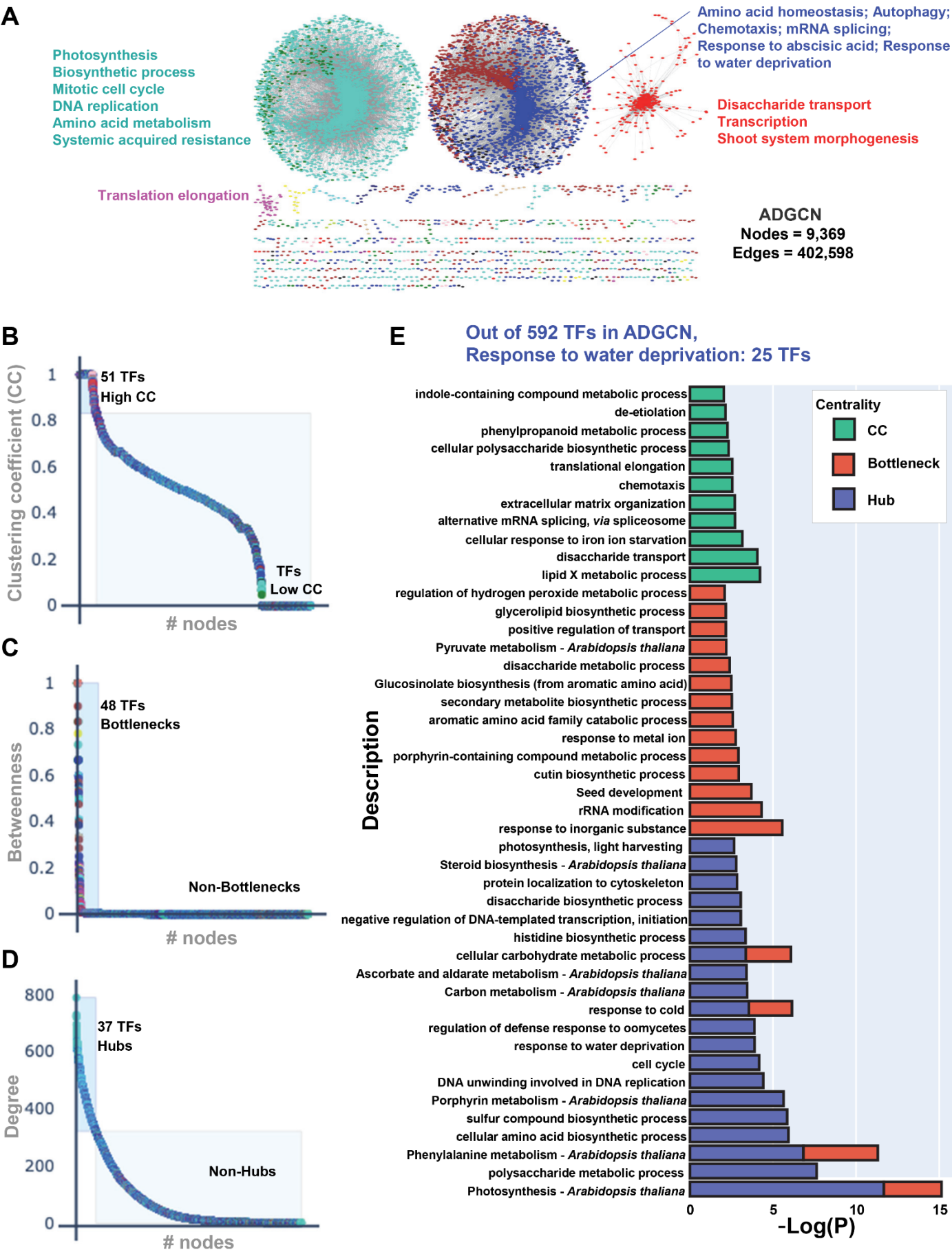


Figure 2. (A) Weighted gene co-expression network analysis (WGCNA) constructed the Arabidopsis Drought-specific Gene Co-expression Network (ADGCN) encompassing 9,370 nodes connected by 402,598 weighted edges (≥ 0.75). Nodes are colored based on their specific module assignment. Most significant modules based on connectivity and clustering coefficient were annotated by GO analysis and the enriched pathways are listed next to their respective modules ($p\text{-value} \leq 0.05$). (B, C, D) ADGCN network analysis plots displaying clustering coefficient (CC) (B), betweenness centrality (C), and degree distribution (D). (E) Functional annotation and pathway enrichment analyses of high CC (green), high bottleneck (red), and high hub (violet) genes.

To identify the major TFs as well as additional transcriptional signatures that can significantly contribute to gene functions in the above-described molecular pathways, we next performed network architectural and centrality analyses. Since the topology of scale-free networks is largely controlled by highly connected nodes, we focused our analyses on clustering coefficient, hubs (highly connected nodes), and bottlenecks (high betweenness nodes). Among the 592 TFs discovered in ADGCN, 51, 48, and 37 TFs exhibited high clustering coefficient, bottlenecks, and hub properties, respectively (Figures 2B, C, D and Table S1). Given that 17 TFs were common between hubs and bottlenecks and only two TFs are shared between bottlenecks and clustering coefficient, cumulatively, we identified 117 non-redundant sets of high-priority TFs that possessed either of the three above-described centrality features. Major families that were enriched in this set of high-priority TFs included ANAC, activator protein 1 (AP1), bHLH, bZIP, Homeobox, MYB, nuclear factor Y (NF-Y), TCP domain protein (TCP), WRKY DNA-binding domain protein (WRKY), and Zinc finger (Table S1). Finally, we also highlighted the functional GO categories of non-TF genes that were classified as high clustering coefficient, bottlenecks, and hub nodes. Notably, “alternative splicing”, “translational elongation” and “polysaccharide biosynthetic process” gene function pathways were enriched in high clustering coefficient nodes (Figure 2B and 2E). The “metabolism” and “amino acid” GO categories were exclusively enriched in bottlenecks (Figure 2C and 2E). Finally, we discovered that hubs were enriched in “cell cycle”, “DNA replication”, “photosynthesis”, “amino acid biosynthetic process” and “response to water deprivation” gene function classifications (Figure 2D and 2E). Overall, the generation of comprehensive co-expression network and network topological analyses discovered potential major master regulators, nodes, and pathways in Arabidopsis drought signaling.

2.3. Dynamic transcriptional regulatory modeling discovered major bifurcated events of the TFs, which participate in the “water deprivation response” pathway

Direct transcriptional gene regulation relies on intricate coordination between TFs and their downstream target genes [32,33]. Therefore, inferring the dynamic regulatory relationships between TFs and their targets can provide a mechanistic view of how multiple TFs collaboratively participate in drought-related transcriptional circuitry at diverse stages of water deficiency stress. Towards this, we employed the iDREM package [32,33] that reconstructed drought-specific reaction networks in an unsupervised manner. Specifically, we integrated the publicly available time-course transcriptome data (GSE76827) with static TF-target datasets and performed dynamic regulatory modeling (see Methods). We discovered significant bifurcation events in which the transcription of a subset of genes deviated from the expression levels of the remainder of the genes over the course of time, as well as identified all the TFs that might govern these sets of genes (Figure 3, Figure S4, Table S2). Furthermore, we also inferred the functional classification of the TFs and their targets based on the GO category data (Figure S4; $p\text{-value} \leq 0.05$). Overall, we showed that 411, 323, 602, 212, and 546 TFs were modeled in diverse paths on day 1, 3, 5, 7 and 9, respectively (Figure 3, Table S2). We subsequently focused our analyses on “response to water deprivation”, “response to stress”, “response to abiotic stimulus”, “response to chemical”, “response to hormone-mediated signal”, and “response to water” gene functional categories. Next, we directed our subsequent analyses on the “response

to water deprivation” GO category. Towards this, we compared all the TFs in this category from our dynamic regulatory modeling with the ADGCN co-expression network and discovered a set of common 25 TFs (Table S2). This set of selected TFs was enriched in bZIP, MYB, NAC, ERF, and AP1 superfamilies. In summary, our dynamic regulatory modeling allowed us to identify the transcriptional paths and major transcriptional events of previously undiscovered 25 TFs of “response to water deprivation” functional classification, which was also one of the categories enriched in the “blue” module of the ADGCN co-expression network (Figure 2A). The identification of these 25 critical TF regulators will substantially broaden our understanding of Arabidopsis drought-related transcriptional circuitry and provide candidate genes for downstream analyses in crops.

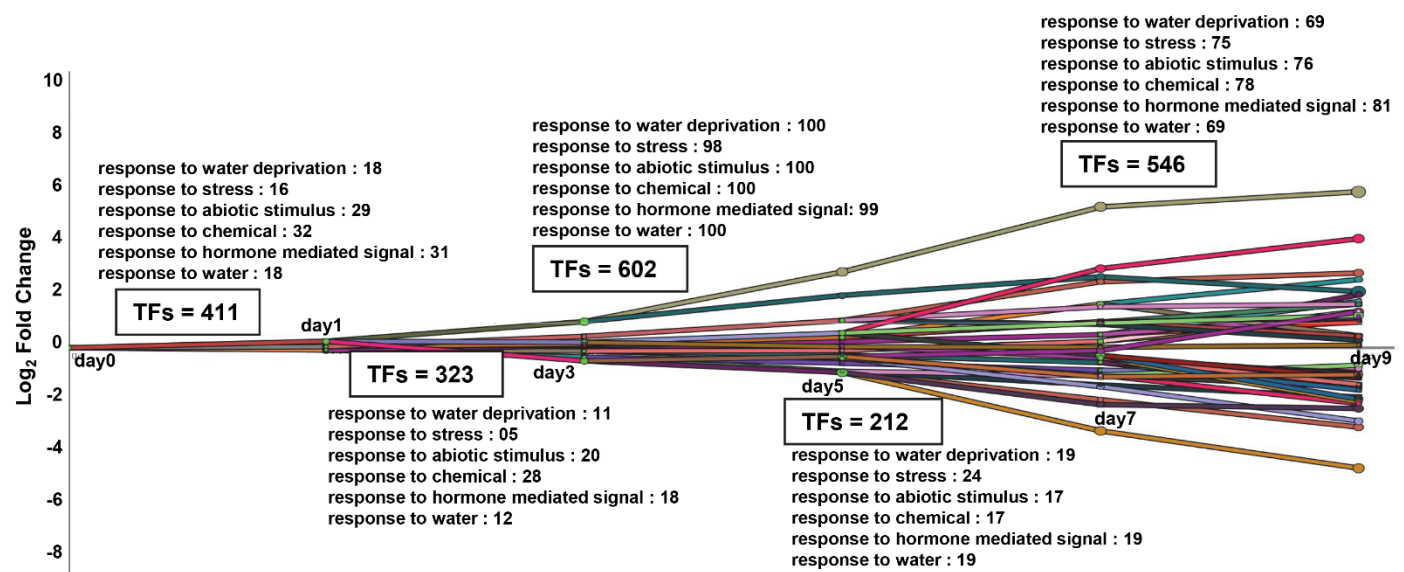


Figure 3. The plot illustrates the interactive Dynamic Regulatory Events Miner (iDREM) of drought response in Arabidopsis from day 0 to 9 at significance $p < 0.01$. The ontology identified the transcription factors (TFs) and genes involved in different functions on different days of drought stress. Each colored line corresponds to a unique transcriptional regulatory path and numbers of TFs representing each functional category are listed in text boxes.

2.4. Mathematical transcriptional simulation of six major TFs acquired by integrating drought-related co-expression data with dynamic transcriptional modeling

To simulate the influence of the regulatory activity of major TFs on their target genes, we used SQUAD simulations on an integrated network that included ADGCN and TF-target association data. For all the downstream simulations and experimental validations, we selected six major TFs that were present among the above-described 25 TFs of “response to water deprivation” and exhibited the properties of high clustering coefficient, hubs, and bottlenecks. This includes AT4G14770 (TESMIN/TSO1-LIKE CXC 2; TCX2/SOL2), AT1G51140 (ABA-RESPONSIVE KINASE SUBSTRATE 1; AKS1/FLOWERING BHLH 3; FBH3), AT5G56840 (a MYB-like TF), AT2G36270 (ABA INSENSITIVE 5; ABI5), AT3G09600 (LHY-CCA1-LIKE5; LCL5/REVEILLE 8; RVE8), and AT2G46590 (DOF AFFECTING GERMINATION 2; DAG2) (Figure 4, Table S2). Interestingly, our dynamic regulatory modeling indicated major bifurcation events for these TFs at day 3, day 5, and day 7 (Table S2). We also showed that TCX2, AKS1/FBH3, AT5G56840, ABI5, LCL5/RVE8, and DAG2 co-regulate with 63, 146, 6, 121, 28, and 2 target genes, respectively. For mathematical simulations, the activation status of these six TFs was individually changed from 0 to 1 over the course of time to predict the dynamic behavior of their target genes. The transcriptional trajectory plots of the target genes illustrated varied levels of amplitude and intensity of activation for the corresponding putative transcriptional targets (Figure

4, Figure S5, Table S2). Overall, we simulated the transcriptional kinetics of target genes for a selected set of six major TFs to gain additional insights into their regulatory influence during the Arabidopsis drought response.

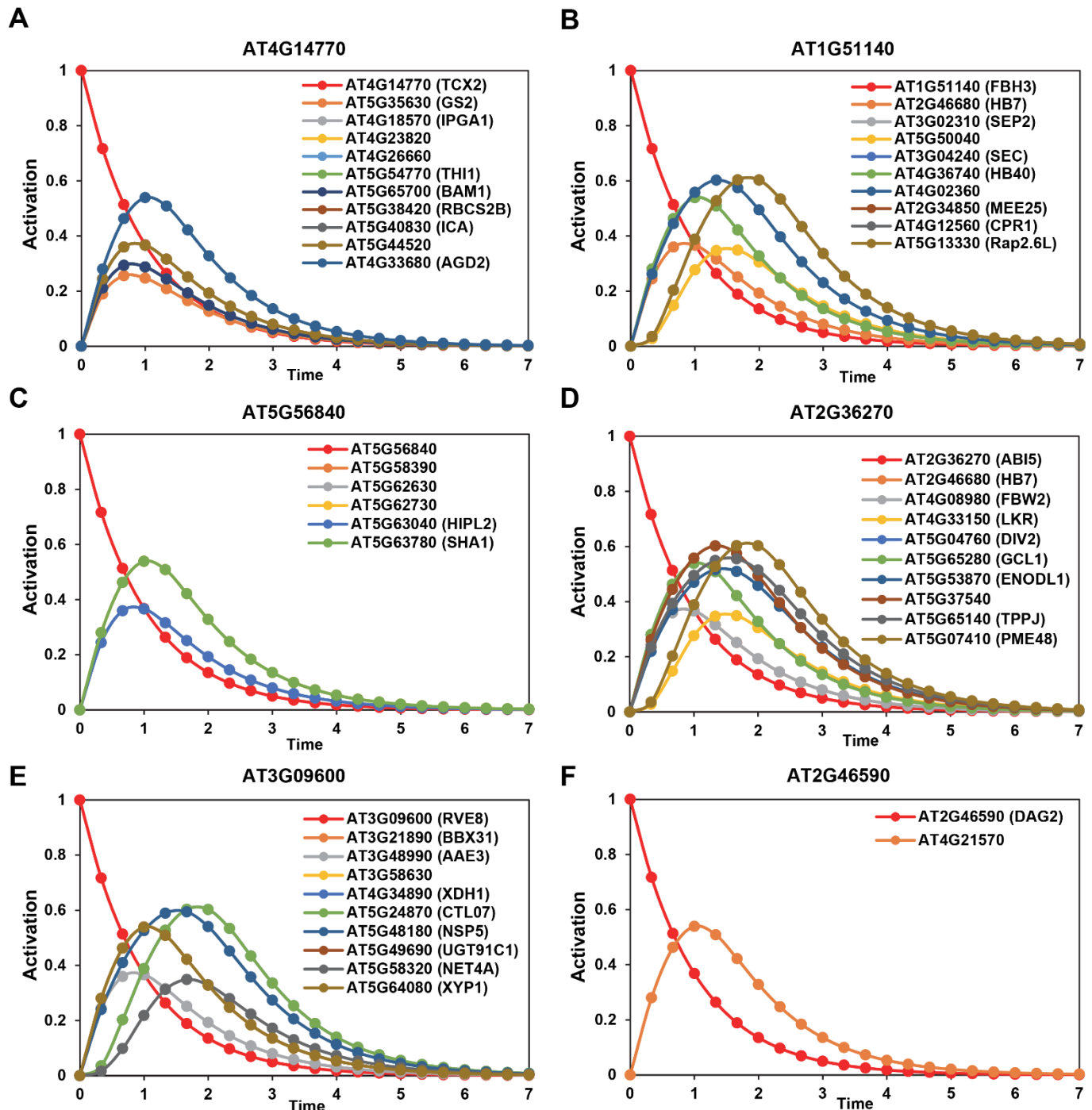


Figure 4. (A–F) The dynamic modeling and simulation of transcription factor regulatory network's activity upon activation of six transcription factors (AT4G14770, AT1G51140, AT5G56840, AT2G36270, AT3G09600, and AT2G46590) identified by SQUAD. Activation patterns of the TFs are shown in red lines whereas individual target genes' activation is illustrated in different colors.

2.5. Experimental evidence for the regulatory relationship between major drought-related TFs and their target genes

To provide experimental validation of the dynamic gene regulation and transcriptional simulations, we performed extensive time-course qRT-PCR experiments in *Arabidopsis* wild-type Col-0 plants. We collected leaf samples of drought-stressed or corresponding well-watered control plants daily over a period of ten days. Based on the activation status, transcriptional trajectories, and amplitudes of gene regulation, we further selected four TFs (TCX2, AKS1/FBH3, ABI5, and LCL5/RVE8) and four targets for each of these TFs (Table S3). TCX2 is a regulator of cell fate transition that controls divisions by regulating stem cell-type-specific networks [37]. We demonstrated that the mRNA levels of TCX2 are down-regulated on day 2 and day 3, while its transcript levels are increased on day 5 and day 6 (Figure 5A). On the other hand, we observed a relative negative transcriptional influence of this TF on its target gene AT4G15790, particularly on day 1, day 5, and day 6 (Figure 5A). Interestingly, the other three tested target genes (AT4G23820; PGF13, Pectin lyase-like superfamily protein, AT5G38420; DEG24, a member of the Rubisco small subunit, and AT4G33680; AGD2, a diaminopimelate aminotransferase) positively co-expressed with TCX2 on day 5 and day 6 (Figure 5B, C, D, E; Table S3). Interestingly, TCX2 is also shown to regulate stomatal cell lineages, further highlighting its connection to drought responses [37]. Our data indicate the potential roles of this stem-cell expressed TF and its target genes in drought, particularly in the mid-phase of drought stress in guard cells.

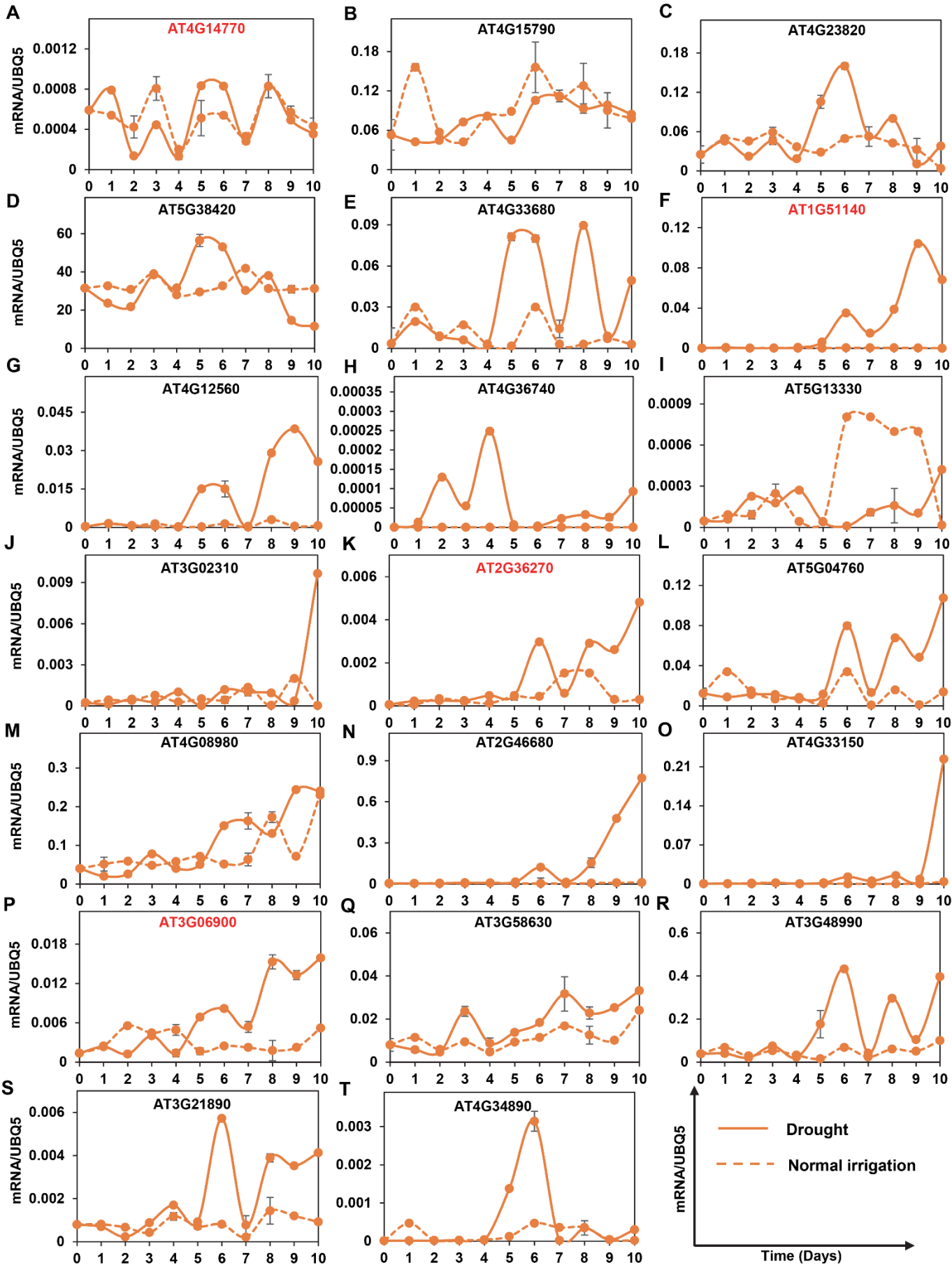


Figure 5. Kinetics of gene expression of transcription factors (red) and their targets in Col-0 under drought (solid lines) and normal irrigation (dotted lines) conditions. RT-qPCR was performed in leaf samples collected daily from day 0 to day 10. Gene expression was assessed using reference gene UBQ5 in four groups: TF1 (A) AT4G14770 and targets (B) AT4G15790, (C) AT4G23820, (D) AT5G38420, (E) AT4G33680; TF2 (F) AT1G51140 and targets (G) AT4G12560, (H) AT4G36740, (I) AT5G13330, (J) AT3G02310; TF3 (K) AT2G36270 and targets (L) AT5G04760, (M) AT4G08980, (N) AT2G46680, (O) AT4G33150; TF4 (P) AT3G06900 and targets (Q) AT3G58630, (R) AT3G48990, (S) AT3G21890, (T) AT4G34890. The graphs represent the mean with standard errors of three technical replicates. Experiments were performed in three biological replications.

AKS1 (also known as FBH3) was shown to be released from the ABA signaling core complex, possibly in the guard cells [38]. Moreover, this basic helix-loop-helix type TF is also involved in photoperiodism flowering [38] (Table S3). Our time-course expression data indicate that AKS1 is induced on day 5 of drought stress and its mRNA accumulation continues to increase for the remainder of the tested drought stress period (Figure 5F, Table S3). While AT4G12560 (CPR1, a defense-responsive gene) positively co-expressed with AKS1 (Figure 5G, Table S3), the transcript abundance of AT4G36740 (HB40, a homeodomain leucine zipper class I (HD-Zip I) protein) was negatively affected by AKS1 (Figure 5H). Moreover, we also showed that the expression levels of AT5G13330 (RAP2.6L, a member of the ERF) were suppressed at the later phase of drought stress (Figure 5I, Table S3). Finally, the transcript levels of AT3G02310 (AGL4, a MADS-box protein) were only heightened on day 10 of drought stress (Figure 5J, Table S3). These data further highlighted the complex involvement of the guard cells in various downstream responses to drought, and shed new light on how a master regulator controls additional TFs at various stages of drought stress.

The third selected gene, AT2G36270 (ABI5), encodes a member of the basic leucine zipper TF family that is involved in ABA signaling during seed maturation and germination [39]. The transcript levels of this master regulator were increased on day 6, day 8, day 9, and day 10 (Figure 5K). Interestingly, two of its target genes (AT5G04760; DIV2 an R-R-type MYB protein, which plays negative roles in salt stress and is required for ABA signaling in Arabidopsis, and AT4G08980; FBW2, a novel negative regulator of AGO1 protein levels that may play a role in ABA signaling) positively co-expressed with ABI5 (Figure 5L, M Table S3). On the other hand, the two additional tested targets of ABI5 (AT2G46680; HB-7 that encodes a putative TF in drought response, and AT4G33150;

LKR that encodes a bifunctional polypeptide lysine-ketoglutarate reductase and saccharopine dehydrogenase involved in lysine degradation; Table S3) were only induced at the very late phase of drought stress (Figure 5N, O). These data further underscore the complex and dynamic roles of ABA signaling at various stages of drought stress and may help provide a mechanistic understanding of how ABI5 controls global transcriptomic changes during drought.

Finally, we examined the expression levels of AT3G09600 (LCL5/RVE8), which encodes an MYB-like TF. LCL5/REV8 was shown to be involved in the regulation of circadian signaling as well as heat shock response [40]. We demonstrated that LCL5/RVE8 expression levels started to rise at day 5 and progressively increased until day 10 of drought stress (Figure 5P). Two of its targets (AT3G58630, a novel TF, and AT3G48990; AAE3, which encodes an oxalyl-CoA synthetase and is required for oxalate degradation) were positively co-expressed with LCL5/RVE8 (Figure 5Q, R Table S3). However, the other two tested targets of LCL5/RVE8 (AT3G21890; BBX31, a B-box type zinc finger family protein, and AT4G34890; XDH1, which encodes a xanthine dehydrogenase that is involved in purine catabolism) were only regulated during day 5 and day 7 of drought stress (Figure 5S, T; Table S3), further highlighting the LCL5/RVE8-mediated crosstalk between heat stress and drought. Therefore, LCL5/RVE8 could constitute a key regulatory node for genetic interventions in the face of the increasingly common heat wave-drought events worldwide.

Taken together, we experimentally characterized the transcriptional regulation patterns of four master TF and their putative targets over a time course of drought stress in Arabidopsis plants.

3. Discussion

Drought is a primary abiotic environmental stress that limits agricultural productivity by impacting plant growth and development during all phases of their lives. Furthermore, global climate change and potential water constraints will likely exacerbate the overall crop yield crisis [41]. Therefore, it is imperative to identify potential novel master regulators that can be used in genome engineering and editing programs to improve plant drought stress tolerance. Here, we used a systems-level approach to reveal phase-specific transcriptional signatures and identify biological pathways pertinent to the Arabidopsis drought transcriptome landscape. Our network science-based analyses unveiled major transcription factors that regulate drought response at various time points. Furthermore, the dynamic regulatory modeling allowed us to uncover major regulatory events and the master regulators responsible for them. Finally, our mathematical transcriptional simulations followed by experimental validation provided potential molecular mechanisms on how master transcriptional regulators operate in an intricate regulatory network to respond at various phases of drought stress.

In our global co-expression network analyses, we found 117 non-redundant sets of high-priority TFs that possess either of three centrality features i.e. clustering coefficient, hubs, and bottlenecks (Figure 2). Since the past decade, network analyses and centrality features have been widely applied to identify the most influential nodes in diverse networks including protein-protein interactions and co-expression networks [14,18,26–29,42,43]. For instance, in a rather focused study, a protein phosphatase 2A (PP2A) network in Arabidopsis was generated [14]. Using closeness centrality score (CCS), 13 PP2A genes were identified that were implicated in the synthesis of amino acids, nucleotides, fatty acids, phytohormones, and vitamin pathways [14]. Intriguingly, our phase-specific unique transcriptome signature analyses discovered that day 5 is enriched with “pyrimidine nucleoside metabolic process”, “regulation of cell communication”, “sulfur and amino acid metabolic” and “ion transport” GO categories (Figure S2) suggesting the potential roles of PP2A in the mid-phase of drought stress in Arabidopsis. In another report, it was demonstrated that Arabidopsis PP2CA is also targeted by the bacterial pathogen *Pseudomonas syringae* to increase plant susceptibility, further supporting the complex role of PP2A genes in various plant stress responses [44].

In maize, over 800 hubs were revealed in a drought-associated co-expression network [45]. The pathway analysis of the non-TFs hubs, glycolysis, and gluconeogenesis pathways was enriched, indicating that they might play an important role during leaf development. Moreover, carbon fixing, nitrogen metabolism, sulfur metabolism, linoleic acid metabolism, monoterpene biosynthesis, and glutathione metabolism were significantly altered during maize stress in drought [45]. While our 117 high-priority TFs were enriched in “alternative splicing”, “translational elongation” and “polysaccharide biosynthetic process” gene function pathways (Figure 2B), our network topology analyses unveiled that the “metabolism” and “amino acid” GO categories are exclusively enriched in bottlenecks (Figure 2C). Finally, we discovered that hubs are enriched in “cell cycle”, “DNA replication”, “photosynthesis”, “amino acid biosynthetic process” and “response to water deprivation” gene function classifications (Figure 2D). The major families of TFs that regulate the above-described pathways belong to ANAC, AP1, bHLH, bZIP, Homeobox, MYB, NF-YA, TCP, WRKY, and Zinc finger (Table S1). Similar functional categories were found in the maize co-expression network, where 49 hub TFs representing bHLH, bZIP, C2H2, Dof, MYB, ERF, and NAC were revealed [45]. Interestingly, 44 of these major TFs were also shown to be involved in four distinct developmental stages of maize [45]. Likewise, the core component of the rice drought co-expression network is enriched in “sequence-specific DNA binding”, “porphyrin and chlorophyll metabolism” [46] that is potentially

regulated by WRKY, Homeodomain, MYB, bZIP, Zinc finger domain, and C2H2 TFs. Overall, our data in Arabidopsis and other studies in monocot and dicot plants indicate the master regulators that are involved in drought-related biological pathways are conserved across diverse plant species. Given the involvement of these TFs in other abiotic and biotic stresses as well as diverse plant development processes, it is highly likely that plants rewire the flow of regulatory information and prioritize growth over survival response during water deficiency conditions, similar to what was reported during biotic stress [47]. Finally, these datasets also point towards global crosstalk of diverse environmental stresses that require major nodes of convergence, possibly master transcriptional regulators. Therefore, the identification of such major transcriptional players is imperative to better inform future breeding programs.

From the integration of the dynamic regulatory network and co-expression network analyses, we focused our mathematical transcriptional simulations and downstream analyses on a set of major TFs that represent “response to water deprivation”-related biological pathways (Figures 2,3,4). Such integrative systems biology-based framework can be expanded to all 117 major TFs, which exemplified other biological GO terms (Figure 2). Our network science-based analysis also signifies the importance of our study to discover and characterize novel players in the drought stress that have been implicated in response to diverse biotic and abiotic stresses, plant developmental processes, and phytohormonal-mediated crosstalk. One such master regulator that was discovered in this study is TCX2, a stem-cell-ubiquitous gene (Figure 5A). Interestingly, TCX2 is conserved in both plants and animals and belongs to the family of CHC proteins that are components of the DREAM cell-cycle regulatory complex [48]. As we proposed its role as a master regulator, recently TCX2 was shown to be a master regulator that orchestrates coordinated stem cell divisions through stem-cell-specific regulatory cascades in roots[37]. Interestingly, TCX2 and its closest homolog SOL1 regulate fate transition and cell divisions in the Arabidopsis stomatal lineage [49], suggesting a multifaceted role of this master regulator in roots and guard cells. While roots play an enormous role in response to drought stress, it's also important to note that regulation of guard cells to avoid water transpiration is equally significant. Thus, our data suggest a critical role of TCX2 in the early- and mid-phase of drought stress (Figure 5A). While TCX2 regulates a suite of diverse genes, we showed that it co-expresses and potentially regulates a pectin lyase (PGF13) [50], a Rubisco subunit (DEG24) [51], and diaminopimelate aminotransferase, which is required for disease resistance to a bacterial pathogen [52]. Cumulatively, the evidence suggests the potential roles of TCX2 in photosynthesis, cell wall metabolism, and disease resistance. Notably, bacterial pathogens use stomata as the main route of entry to the leaf cells [53–55] further highlighting the roles of this major regulator in the crosstalk of biotic and abiotic stresses with plant developmental processes.

Considering that stomatal closure is regulated by ABA to avoid further water loss during drought stress, it was unsurprising that two additional master regulators identified in the current study are major components of the ABA signaling pathway (Figure 5). AKS1 was shown to be released from the ABA signaling core complex, possibly in the guard cells, and connects between ABA core machinery and ABA-responsive genes [38]. Interestingly, our dynamic modeling, transcriptional simulations, and gene expression study further connect AKS1 to the phytohormone crosstalk as it co-expresses and potentially regulates an Ethylene Responsive Factor (ERF) transcription factor, RAP2.6L (Figure 5I). While RAP2.6L was previously shown to be involved in resistance against a bacterial pathogen [56], AKS1 also potentially regulates CPR1 (a salicylic acid-dependent defense gene) (Figure 5G). Collectively, these data further underscore the importance of AKS1-dependent regulation of ABA-ET-SA crosstalk and potentially the control of stomatal aperture. ABI5 is the second major component of ABA signaling that is required at various stages of plant development, drought stress, and immune responses [39,44]. Among other targets, we experimentally demonstrated that ABI5 potentially regulates two ABA signaling players DIV2 and FBW2 (Figure 5L, M) that negatively regulates ABA signaling

[57,58]. Interestingly, we also showed the co-regulation of ABI5 with HB-7 (Figure 5N) and AKS1 with HB-40 (Figure 5H). Both of these homeodomain leucine zipper class I (HD-Zip I) possessed the properties of hubs and bottlenecks in our analyses, indicating the importance of these two key players in ADGCN (Table S1). Remarkably, apple plants overexpressing HB-7 were more tolerant to drought treatment, maintained high rate of photosynthesis and accumulated increased biomass suggesting the potential use of our network science-based discovery in genome editing and engineering programs for economically important plant species [59,60].

In summary, our systems biology-fueled and network science-driven approach to model the dynamic transcriptional regulation in drought stress discovered novel master regulators that are involved in diverse biological pathways. Our integrative framework also highlights the intertwined relationships among phytohormone signaling networks, biotic/abiotic stresses, stomatal regulation/guard cell formation, and diverse plant developmental processes.

4. Materials and Methods

4.1. Plant growth conditions

All work was performed with *Arabidopsis thaliana* ecotype Col-0. Col-0 seeds were purchased from the Arabidopsis Biological Resource Center, directly sowed on Fafard Germination Mix in 90 mm × 68 mm Pöppelmann TEKUR® round pots, and cold stratified in the cold room at 4°C for three days. Plants were grown at 23 °C with a 12 h light/12 h dark cycle with a light intensity of 120 $\mu\text{E}/\text{m}^2/\text{s}$ and watered every three days with tap water to keep soil moisture. Ten-day-old seedlings were transplanted to 12 × 6 Landmark Plastic flats with continuous irrigation till day 18. Half of the 18-day-old population was grown without further input of water, while the other half was supplied with 3 ml tap water per plant 2 h prior to sampling. Samples were collected daily from day 18 to day 28 and flash-frozen in liquid nitrogen followed by storing in -80 °C. Three leaves were harvested for each replicate and three biological replicates were prepared.

4.2. RNA extractions and quantitative real-time polymerase chain reaction (qRT-PCR)

Leaf samples of Col-0 were ground with Bead Ruptor 96-Well Plate Homogenizer. Total RNA was extracted using TRIzol (Invitrogen) according to the manufacturer's protocol. Briefly, 1ml TRIzol was used per sample. The RNA pellets were dissolved in 20 μl DEPC-treated water and quantified by BioPhotometer Plus (Eppendorf). 10 μg of RNA samples were then treated with DNase using the TURBO DNA-free™ Kit (Ambion) in a 20 μl reaction. cDNA was synthesized using SuperScript IV reverse transcriptase first-strand synthesis kit (Invitrogen) with 2 μg DNA-free RNA in a 10 μl reaction as described previously [61]. PCR programs from both DNase treatment and reverse transcription reaction was performed on Applied Biosystems 96-Well Thermal Cycler. qRT-PCR was performed on an ABI 7500 Fast PCR System (ThermoFisher Scientific) with 2X PowerUp SYBR green master mix (Applied Biosystems, ThermoFisher Scientific) using the following settings: 50 °C for 2 min and 95 °C for 10 min followed by 40 cycles of 95 °C for 15 sec, 55 °C for 15 sec and 72 °C for 1 min. Primer sequences are listed in Table S3.

4.3. Drought stress analysis data acquisition

Drought-related data pertinent to this study were downloaded from NCBI Gene Expression Omnibus (GEO) database [62]. Specifically, we employed the GSE76827 dataset that consists of transcriptome pertinent to drought-stressed aerial parts of *Arabidopsis* at 0, 1, 3, 5, 7, and 9 days. A custom Python script (version 3.8) was used to preprocess the sample data to filter out low variance genes. During the preprocessing, genes that had zero expression values for more than 30 percent of the samples were pruned to maintain adequately high variance, which is necessary for the co-expression network construction.

This approach provided data that is ready to proceed for downstream network-based and additional analyses.

4.4. Differentially expressed genes (DEGs) analysis and functional annotation pathway analyses

The processed transcriptome data from GSE76827 was subjected to iDEP[63], which is an interactive DEG (iDEP) analysis web tool, to compute DEGs between drought and mock treatments at diverse time points. Briefly, iDEP utilizes DESeq2 R package to analyze DEGs [64]. We used a threshold of 1.5 fold change (FC) which is equivalent to 0.58 log₂ Fold Change (log₂FC) and a false discovery rate (FDR) ≤ 0.05 was set for DEGs [21]. DEGs were further subjected to hierarchical clustering. Functional annotation and pathway enrichment analyses were performed using Metascape using the cutoff 1.30 in -log₁₀ scale (p-value ≤ 0.05) [35].

4.5. Weighted Gene Co-expression Network Analysis (WGCNA) and network centrality analyses

Towards generating an Arabidopsis Drought-specific Gene Co-expression Network (ADGCN), we utilized the GSE76827 dataset that consists of transcriptome pertinent to drought stressed aerial parts of Arabidopsis at 0, 1, 3, 5, 7, and 9 days. We employed the WGCNA and (optional Limma) R package, which provides a wide spectrum of R functions and carries out diverse weighted correlation network analyses. Specifically, this package includes functions for the construction of a co-expression network, detection of various modules with different correlation coefficients, analysis of a wide range of network topological features, and simulation as well as visualization of data and network [36]. WGCNA follows three major steps to build a co-expression network: This includes: 1) defining an adjacency matrix among all the genes based on pair-wise correlations, 2) calculation of network centrality features from the adjacency matrix and converted into a dissimilarity measure, and 3) application of hierarchical clustering by employing this dissimilarity measure. This is followed by tree cutting using either a static or a dynamic height cut. This analysis results in clusters or modules of genes with significance [36]. The data is formed into modules based on correlation patterns, colors are assigned to differentiate the modules. The nodes and edges data will be used to perform network analysis in the next step [42]. In ADGCN co-expression network, the nodes correspond to genes, while weighted edges (≥ 0.75) define the interconnectedness based on expression correlations [65]. The edges-nodes data and module assignment data were further imported to visualize networks with module-assigned colors in Cytoscape [66]. After network construction, the topological features such as eigenvector centrality, degree centrality, betweenness centrality, closeness centrality, and information centrality are computed. After network construction, the topological features such as degree centrality, betweenness centrality, and clustering coefficient were computed using the Python-based NetworkX 2.6 package [67].

4.6. Dynamic gene regulatory event mining and dynamic simulations.

To explore the dynamic behavior of water stress on Arabidopsis, we performed the Interactive Dynamic Regulatory Events Miner (iDREM) analysis on the drought-specific transcription response in Arabidopsis. iDREM software employs static protein-DNA interaction data including Chromatin Immunoprecipitation Sequencing (ChIP-Seq) ChIP-seq etc. and integrates with time-series gene expression data to simulate dynamic regulatory networks [32]. For TF-target dataset, we curated TF and target gene regulatory network by including vast datasets from Arabidopsis thaliana Regulatory Network (AtRegNet) [68], Plant Cistrome Database (DAP_seq) [69], Arabidopsis transcriptional regulatory map (ARTM) [69], Curated_1 [70], TF2Network (Curated_2) [71] and Ath [68,70–72]. The Arabidopsis Information Resource (TAIR) [73], TF-gene interaction datasets, the European Bioinformatics Institute (EBI)-based functional ontology [74], and drought-specific Arabidopsis expression data were used to perform iDREM analysis. iDREM highlighted

major bifurcation events and displayed TFs that exhibit significant feature changes at a given time point. The score threshold was kept as 0.01 and significance based on path significance conditional on a split, where the minimum split percentage was 50%. On each day of bifurcation, we highlighted the overall TFs expressed, afterward selected the top 100 TFs from the list based on a highly regulated quantity of genes, and compared them with the Arabidopsis co-expression network. GO term categories were selected with a p-value ≤ 0.05 . Finally, 6 TFs were selected from this group based on high network centrality features including degree, betweenness, and clustering coefficient. Next, we focused on “water deprivation” functional annotation. Mathematical simulations were performed using SQUAD (Standardized Qualitative Dynamical systems) [34] as described before [21].

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1.

Figure S1: The pipeline describing the discovery of novel master regulators in drought.

Figure S2. The gene ontology enrichment analysis of Arabidopsis unique DEGs on each (Day 3 and Day 5) after drought stress (p-value ≤ 0.05).

Figure S3. (A, B) The weighted gene co-expression network analysis (WGCNA) parameters for Arabidopsis drought transcriptome analysis and module identification. (C, D) The violin plot of clustering coefficient and connectivity to identify the most significant modules of ADGCN based on the cutoff of an average value of 0.43 and 144 for clustering coefficient and connectivity, respectively.

Figure S4. The iDREM analysis identified significantly enriched pathways and several TFs associated with the regulation of gene profiles in those pathways.

Figure S5. (A-D) The SQUAD simulation of gene activity upon the activation of TFs (AT4G14770, AT1G51140, AT2G36270, AT3G09600) in drought stress.

Table S 1: Gene expression, co-expression and network centrality analyses

Table S 2: iDREM and SQUAD analyses

Table S3: Arabidopsis transcription factors and their targets included in experimental analyses; PCR primers list

Author Contributions: N.K. and B.K.M. performed the computational analyses. J.L. and B.M. performed qRT-PCR experiments. N.K., B.K.M., and J.L. analyzed the data. K.P.M. and M.S.M. coordinated the research program, oversaw the experimental work, and wrote the manuscript. All authors discussed the results, critically reviewed the manuscript, and provided feedback.

Funding: This work was supported by the National Science Foundation awards IOS-1557796 and IOS-2038872 to M.S.M., and IOS-1350244 and IOS-2038872 to K.P.M.

Institutional Review Board Statement: “Not applicable”

Informed Consent Statement: “Not applicable.”

Data Availability Statement: The dataset analyzed in this study (GSE76827) can be found in the NCBI Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>).

Acknowledgments: We are grateful to Mr. Adeel Mehmood, Mr. Revanth K. Nalluri and Ms. Mi-kayla Jackson for technical assistance in bioinformatics analyses and laboratory work.

Conflicts of Interest: The authors declare no conflict of interest

References

1. Rey, D.; Holman, I. P.; Knox, J. W., Developing drought resilience in irrigated agriculture in the face of increasing water scarcity. *Reg Environ Change* **2017**, 17, (5), 1527-1540.

2. Fahad, S.; Bajwa, A. A.; Nazir, U.; Anjum, S. A.; Farooq, A.; Zohaib, A.; Sadia, S.; Nasim, W.; Adkins, S.; Saud, S.; Ihsan, M. Z.; Alharby, H.; Wu, C.; Wang, D.; Huang, J., Crop Production under Drought and Heat Stress: Plant Responses and Management Options. *Front Plant Sci* **2017**, *8*, 1147.
3. Leng, G.; Hall, J., Crop yield sensitivity of global major agricultural countries to droughts and the projected changes in the future. *Sci Total Environ* **2019**, *654*, 811-821.
4. Obidiegwu, J. E.; Bryan, G. J.; Jones, H. G.; Prashar, A., Coping with drought: stress and adaptive responses in potato and perspectives for improvement. *Front Plant Sci* **2015**, *6*, 542.
5. Lv, L.; Zhang, W.; Sun, L.; Zhao, A.; Zhang, Y.; Wang, L.; Liu, Y.; Li, Z.; Li, H.; Chen, X., Gene co-expression network analysis to identify critical modules and candidate genes of drought-resistance in wheat. *PLoS One* **2020**, *15*, (8), e0236186.
6. Tolosa, L. N.; Zhang, Z., The Role of Major Transcription Factors in Solanaceous Food Crops under Different Stress Conditions: Current and Future Perspectives. *Plants (Basel)* **2020**, *9*, (1).
7. Ximenez-Embun, M. G.; Gonzalez-Guzman, M.; Arbona, V.; Gomez-Cadenas, A.; Ortego, F.; Castanera, P., Plant-Mediated Effects of Water Deficit on the Performance of Tetranychus evansi on Tomato Drought-Adapted Accessions. *Front Plant Sci* **2018**, *9*, 1490.
8. Bista, D. R.; Heckathorn, S. A.; Jayawardena, D. M.; Mishra, S.; Boldt, J. K., Effects of Drought on Nutrient Uptake and the Levels of Nutrient-Uptake Proteins in Roots of Drought-Sensitive and -Tolerant Grasses. *Plants (Basel)* **2018**, *7*, (2).
9. Bouzid, M.; He, F.; Schmitz, G.; Hausler, R. E.; Weber, A. P. M.; Mettler-Altmann, T.; De Meaux, J., Arabidopsis species deploy distinct strategies to cope with drought stress. *Ann Bot* **2019**, *124*, (1), 27-40.
10. Gamboa-Tuz, S. D.; Pereira-Santana, A.; Zamora-Briseno, J. A.; Castano, E.; Espadas-Gil, F.; Ayala-Sumuano, J. T.; Keb-Llanes, M. A.; Sanchez-Teyer, F.; Rodriguez-Zapata, L. C., Transcriptomics and co-expression networks reveal tissue-specific responses and regulatory hubs under mild and severe drought in papaya (Carica papaya L.). *Sci Rep* **2018**, *8*, (1), 14539.
11. Chen, M.; Zhu, X.; Zhang, Y.; Du, Z.; Chen, X.; Kong, X.; Sun, W.; Chen, C., Drought stress modify cuticle of tender tea leaf and mature leaf for transpiration barrier enhancement through common and distinct modes. *Sci Rep* **2020**, *10*, (1), 6696.
12. Khadka, V. S.; Vaughn, K.; Xie, J.; Swaminathan, P.; Ma, Q.; Cramer, G. R.; Fennell, A. Y., Transcriptomic response is more sensitive to water deficit in shoots than roots of Vitis riparia (Michx.). *BMC Plant Biol* **2019**, *19*, (1), 72.
13. Bian, Z.; Wang, Y.; Zhang, X.; Grundy, S.; Hardy, K.; Yang, Q.; Lu, C., A Transcriptome Analysis Revealing the New Insight of Green Light on Tomato Plant Growth and Drought Stress Tolerance. *Front Plant Sci* **2021**, *12*, 649283.
14. Khan, Z. H.; Agarwal, S.; Rai, A.; Memaya, M. B.; Mehrotra, S.; Mehrotra, R., Co-expression network analysis of protein phosphatase 2A (PP2A) genes with stress-responsive genes in Arabidopsis thaliana reveals 13 key regulators. *Sci Rep* **2020**, *10*, (1), 21480.
15. Nakashima, K.; Yamaguchi-Shinozaki, K.; Shinozaki, K., The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. *Front Plant Sci* **2014**, *5*, 170.
16. Baldoni, E.; Genga, A.; Cominelli, E., Plant MYB Transcription Factors: Their Role in Drought Response Mechanisms. *Int J Mol Sci* **2015**, *16*, (7), 15811-51.
17. Singh, D.; Laxmi, A., Transcriptional regulation of drought response: a tortuous network of transcriptional factors. *Front Plant Sci* **2015**, *6*, 895.
18. Tully, J. P.; Hill, A. E.; Ahmed, H. M.; Whitley, R.; Skjellum, A.; Mukhtar, M. S., Expression-based network biology identifies immune-related functional modules involved in plant defense. *BMC Genomics* **2014**, *15*, 421.
19. Mishra, B.; Sun, Y.; Ahmed, H.; Liu, X.; Mukhtar, M. S., Global temporal dynamic landscape of pathogen-mediated subversion of Arabidopsis innate immunity. *Sci Rep* **2017**, *7*, (1), 7849.
20. Alvarez, J. M.; Brooks, M. D.; Swift, J.; Coruzzi, G. M., Time-Based Systems Biology Approaches to Capture and Model Dynamic Gene Regulatory Networks. *Annu Rev Plant Biol* **2021**, *72*, 105-131.
21. Mishra, B.; Sun, Y.; Howton, T. C.; Kumar, N.; Mukhtar, M. S., Dynamic modeling of transcriptional gene regulatory network uncovers distinct pathways during the onset of Arabidopsis leaf senescence. *NPJ Syst Biol Appl* **2018**, *4*, 35.
22. Ahmed, H.; Howton, T. C.; Sun, Y.; Weinberger, N.; Belkhadir, Y.; Mukhtar, M. S., Network biology discovers pathogen contact points in host protein-protein interactomes. *Nat Commun* **2018**, *9*, (1), 2312.
23. Mishra, B.; Kumar, N.; Mukhtar, M. S., Network biology to uncover functional and structural properties of the plant immune system. *Curr Opin Plant Biol* **2021**, *62*, 102057.
24. Mishra, B.; Kumar, N.; Mukhtar, M. S., Systems Biology and Machine Learning in Plant-Pathogen Interactions. *Mol Plant Microbe Interact* **2019**, *32*, (1), 45-55.
25. Garbutt, C. C.; Bangalore, P. V.; Kannar, P.; Mukhtar, M. S., Getting to the edge: protein dynamical networks as a new frontier in plant-microbe interactions. *Front Plant Sci* **2014**, *5*, 312.
26. Naqvi, R. Z.; Zaidi, S. S.; Akhtar, K. P.; Strickler, S.; Woldemariam, M.; Mishra, B.; Mukhtar, M. S.; Scheffler, B. E.; Scheffler, J. A.; Jander, G.; Mueller, L. A.; Asif, M.; Mansoor, S., Transcriptomics reveals multiple resistance mechanisms against cotton leaf curl disease in a naturally immune cotton species, Gossypium arboreum. *Sci Rep* **2017**, *7*, (1), 15880.
27. Zaidi, S. S.; Naqvi, R. Z.; Asif, M.; Strickler, S.; Shakir, S.; Shafiq, M.; Khan, A. M.; Amin, I.; Mishra, B.; Mukhtar, M. S.; Scheffler, B. E.; Scheffler, J. A.; Mueller, L. A.; Mansoor, S., Molecular insight into cotton leaf curl geminivirus disease resistance in cultivated cotton (Gossypium hirsutum). *Plant Biotechnol J* **2020**, *18*, (3), 691-706.

28. Naqvi, R. Z.; Zaidi, S. S.; Mukhtar, M. S.; Amin, I.; Mishra, B.; Strickler, S.; Mueller, L. A.; Asif, M.; Mansoor, S., Transcriptomic analysis of cultivated cotton *Gossypium hirsutum* provides insights into host responses upon whitefly-mediated transmission of cotton leaf curl disease. *PLoS One* **2019**, *14*, (2), e0210011.
29. Mishra, B.; Athar, M.; Mukhtar, M. S., Transcriptional circuitry atlas of genetic diverse unstimulated murine and human macrophages define disparity in population-wide innate immunity. *Sci Rep* **2021**, *11*, (1), 7373.
30. Bizouerne, E.; Buitink, J.; Vu, B. L.; Vu, J. L.; Esteban, E.; Pasha, A.; Provart, N.; Verdier, J.; Leprince, O., Gene co-expression analysis of tomato seed maturation reveals tissue-specific regulatory networks and hubs associated with the acquisition of desiccation tolerance and seed vigour. *BMC Plant Biol* **2021**, *21*, (1), 124.
31. Sakuraba, Y.; Kim, Y. S.; Han, S. H.; Lee, B. D.; Paek, N. C., The Arabidopsis Transcription Factor NAC016 Promotes Drought Stress Responses by Repressing AREB1 Transcription through a Trifurcate Feed-Forward Regulatory Loop Involving NAP. *Plant Cell* **2015**, *27*, (6), 1771-87.
32. Ding, J.; Hagood, J. S.; Ambalavanan, N.; Kaminski, N.; Bar-Joseph, Z., iDREM: Interactive visualization of dynamic regulatory networks. *PLoS Comput Biol* **2018**, *14*, (3), e1006019.
33. Mishra, B.; Kumar, N.; Liu, J.; Pajeroska-Mukhtar, K. M., Dynamic Regulatory Event Mining by iDREM in Large-Scale Multi-omics Datasets During Biotic and Abiotic Stress in Plants. *Methods Mol Biol* **2021**, 2328, 191-202.
34. Di Cara, A.; Garg, A.; De Micheli, G.; Xenarios, I.; Mendoza, L., Dynamic simulation of regulatory networks using SQUAD. *BMC Bioinformatics* **2007**, *8*, 462.
35. Zhou, Y.; Zhou, B.; Pache, L.; Chang, M.; Khodabakhshi, A. H.; Tanaseichuk, O.; Benner, C.; Chanda, S. K., Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **2019**, *10*, (1), 1523.
36. Langfelder, P.; Horvath, S., WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **2008**, *9*, 559.
37. Clark, N. M.; Buckner, E.; Fisher, A. P.; Nelson, E. C.; Nguyen, T. T.; Simmons, A. R.; de Luis Balaguer, M. A.; Butler-Smith, T.; Sheldon, P. J.; Bergmann, D. C.; Williams, C. M.; Sozzani, R., Stem-cell-ubiquitous genes spatiotemporally coordinate division through regulation of stem-cell-specific gene networks. *Nat Commun* **2019**, *10*, (1), 5574.
38. Takahashi, Y.; Ebisu, Y.; Shimazaki, K. I., Reconstitution of Abscissic Acid Signaling from the Receptor to DNA via bHLH Transcription Factors. *Plant Physiol* **2017**, *174*, (2), 815-822.
39. Skubacz, A.; Daszkowska-Golec, A.; Szarejko, I., The Role and Regulation of ABI5 (ABA-Insensitive 5) in Plant Development, Abiotic Stress Responses and Phytohormone Crosstalk. *Front Plant Sci* **2016**, *7*, 1884.
40. Yang, P.; Wang, J.; Huang, F. Y.; Yang, S.; Wu, K., The Plant Circadian Clock and Chromatin Modifications. *Genes (Basel)* **2018**, *9*, (11).
41. Salehi, M., Global water shortage and potable water safety; Today's concern and tomorrow's crisis. *Environ Int* **2022**, *158*, 106936.
42. Kumar, N.; Mishra, B.; Mehmood, A.; Mohammad, A.; Mukhtar, M. S., Integrative Network Biology Framework Elucidates Molecular Mechanisms of SARS-CoV-2 Pathogenesis. *iScience* **2020**, *23*, (9), 101526.
43. McCormack, M. E.; Lopez, J. A.; Crocker, T. H.; Mukhtar, M. S., Making the right connections: network biology and plant immune system dynamics. *Current Plant Biology* **2016**, *5*, 2-12.
44. Liu, X.; Afrin, T.; Pajeroska-Mukhtar, K. M., Arabidopsis GCN2 kinase contributes to ABA homeostasis and stomatal immunity. *Commun Biol* **2019**, *2*, 302.
45. Miao, Z.; Han, Z.; Zhang, T.; Chen, S.; Ma, C., A systems approach to a spatio-temporal understanding of the drought stress response in maize. *Sci Rep* **2017**, *7*, (1), 6590.
46. Lv, Y.; Xu, L.; Dossa, K.; Zhou, K.; Zhu, M.; Xie, H.; Tang, S.; Yu, Y.; Guo, X.; Zhou, B., Identification of putative drought-responsive genes in rice using gene co-expression analysis. *Bioinformation* **2019**, *15*, (7), 480-489.
47. Pajeroska-Mukhtar, K. M.; Wang, W.; Tada, Y.; Oka, N.; Tucker, C. L.; Fonseca, J. P.; Dong, X., The HSF-like transcription factor TBF1 is a major molecular switch for plant growth-to-defense transition. *Curr Biol* **2012**, *22*, (2), 103-12.
48. Desvoves, B.; Gutierrez, C., Roles of plant retinoblastoma protein: cell cycle and beyond. *EMBO J* **2020**, *39*, (19), e105802.
49. Simmons, A. R.; Davies, K. A.; Wang, W.; Liu, Z.; Bergmann, D. C., SOL1 and SOL2 regulate fate transition and cell divisions in the Arabidopsis stomatal lineage. *Development* **2019**, 146, (3).
50. Ohashi, T.; Sari, N.; Misaki, R.; Fujiyama, K., Biochemical characterization of Arabidopsis clade F polygalacturonase shows a substrate preference toward oligogalacturonic acids. *J Biosci Bioeng* **2022**, *133*, (1), 1-7.
51. Jing, Y.; Shi, L.; Li, X.; Zheng, H.; Gao, J.; Wang, M.; He, L.; Zhang, W., OXS2 is Required for Salt Tolerance Mainly through Associating with Salt Inducible Genes, CA1 and Araport11, in Arabidopsis. *Sci Rep* **2019**, *9*, (1), 20341.
52. Song, J. T.; Lu, H.; Greenberg, J. T., Divergent roles in Arabidopsis thaliana development and defense of two homologous genes, aberrant growth and death2 and AGD2-LIKE DEFENSE RESPONSE PROTEIN1, encoding novel aminotransferases. *Plant Cell* **2004**, *16*, (2), 353-66.
53. Mukhtar, M. S.; McCormack, M. E.; Argueso, C. T.; Pajeroska-Mukhtar, K. M., Pathogen Tactics to Manipulate Plant Cell Death. *Curr Biol* **2016**, *26*, (13), R608-R619.
54. Gonzalez-Fuente, M.; Carrere, S.; Monachello, D.; Marsella, B. G.; Cazale, A. C.; Zischek, C.; Mitra, R. M.; Reze, N.; Cottret, L.; Mukhtar, M. S.; Lurin, C.; Noel, L. D.; Peeters, N., EffectorK, a comprehensive resource to mine for Ralstonia, Xanthomonas, and other published effector interactors in the Arabidopsis proteome. *Mol Plant Pathol* **2020**, *21*, (10), 1257-1270.
55. Sun, Y.; Detchemendy, T. W.; Pajeroska-Mukhtar, K. M.; Mukhtar, M. S., NPR1 in JazzSet with Pathogen Effectors. *Trends Plant Sci* **2018**, *23*, (6), 469-472.

56. Sun, F.; Liu, P.; Xu, J.; Dong, H., Mutation in RAP2.6L, a transactivator of the ERF transcription factor family, enhances Arabidopsis resistance to *Pseudomonas syringae*. *Physiological and Molecular Plant Pathology* **2010**, *74*, (5-6), 295-302.
57. Earley, K.; Smith, M.; Weber, R.; Gregory, B.; Poethig, R., An endogenous F-box protein regulates ARGONAUTE1 in *Arabidopsis thaliana*. *Silence* **2010**, *1*, (1), 15.
58. Fang, Q.; Wang, Q.; Mao, H.; Xu, J.; Wang, Y.; Hu, H.; He, S.; Tu, J.; Cheng, C.; Tian, G.; Wang, X.; Liu, X.; Zhang, C.; Luo, K., AtDIV2, an R-R-type MYB transcription factor of Arabidopsis, negatively regulates salt stress by modulating ABA signaling. *Plant Cell Rep* **2018**, *37*, (11), 1499-1511.
59. Zhao, S.; Wang, H.; Jia, X.; Gao, H.; Mao, K.; Ma, F., The HD-Zip I transcription factor MdHB7-like confers tolerance to salinity in transgenic apple (*Malus domestica*). *Physiol Plant* **2021**, *172*, (3), 1452-1464.
60. Zhao, S.; Gao, H.; Jia, X.; Wei, J.; Mao, K.; Ma, F., MdHB-7 Regulates Water Use Efficiency in Transgenic Apple (*Malus domestica*) Under Long-Term Moderate Water Deficit. *Front Plant Sci* **2021**, *12*, 740492.
61. McCormack, M. E.; Liu, X.; Jordan, M. R.; Pajerowska-Mukhtar, K. M., An improved high-throughput screening assay for tunicamycin sensitivity in Arabidopsis seedlings. *Front Plant Sci* **2015**, *6*, 663.
62. Barrett, T.; Trupin, D. B.; Wilhite, S. E.; Ledoux, P.; Rudnev, D.; Evangelista, C.; Kim, I. F.; Soboleva, A.; Tomashevsky, M.; Marshall, K. A.; Phillippy, K. H.; Sherman, P. M.; Muerter, R. N.; Edgar, R., NCBI GEO: archive for high-throughput functional genomic data. *Nucleic Acids Res* **2009**, *37*, (Database issue), D885-90.
63. Ge, S. X.; Son, E. W.; Yao, R., iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **2018**, *19*, (1), 534.
64. Love, M. I.; Huber, W.; Anders, S., Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **2014**, *15*, (12), 550.
65. Horvath, S.; Dong, J., Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol* **2008**, *4*, (8), e1000117.
66. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T., Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **2003**, *13*, (11), 2498-504.
67. Hagberg, A.; Swart, P.; Schult, D., Exploring network structure, dynamics, and function using NetworkX. **2008**.
68. Palaniswamy, S. K.; James, S.; Sun, H.; Lamb, R. S.; Davuluri, R. V.; Grotewold, E., AGRIS and AtRegNet: a platform to link cis-regulatory elements and transcription factors into regulatory networks. *Plant Physiol* **2006**, *140*, (3), 818-29.
69. O'Malley, R. C.; Huang, S. C.; Song, L.; Lewsey, M. G.; Bartlett, A.; Nery, J. R.; Galli, M.; Gallavotti, A.; Ecker, J. R., Cistrome and Episcistrome Features Shape the Regulatory DNA Landscape. *Cell* **2016**, *166*, (6), 1598.
70. Yu, C. P.; Lin, J. J.; Li, W. H., Positional distribution of transcription factor binding sites in *Arabidopsis thaliana*. *Sci Rep* **2016**, *6*, 25164.
71. Kulkarni, S. R.; Vaneechoutte, D.; Van de Velde, J.; Vandepoele, K., TF2Network: predicting transcription factor regulators and gene regulatory networks in Arabidopsis using publicly available binding site information. *Nucleic Acids Res* **2018**, *46*, (6), e31.
72. Jin, J.; He, K.; Tang, X.; Li, Z.; Lv, L.; Zhao, Y.; Luo, J.; Gao, G., An Arabidopsis Transcriptional Regulatory Map Reveals Distinct Functional and Evolutionary Features of Novel Transcription Factors. *Mol Biol Evol* **2015**, *32*, (7), 1767-73.
73. Rhee, S. Y.; Beavis, W.; Berardini, T. Z.; Chen, G.; Dixon, D.; Doyle, A.; Garcia-Hernandez, M.; Huala, E.; Lander, G.; Montoya, M.; Miller, N.; Mueller, L. A.; Mundodi, S.; Reiser, L.; Tacklind, J.; Weems, D. C.; Wu, Y.; Xu, I.; Yoo, D.; Yoon, J.; Zhang, P., The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Res* **2003**, *31*, (1), 224-8.
74. Cote, R.; Reisinger, F.; Martens, L.; Barsnes, H.; Vizcaino, J. A.; Hermjakob, H., The Ontology Lookup Service: bigger and better. *Nucleic Acids Res* **2010**, *38*, (Web Server issue), W155-60.