



Transcriptional Network Analysis Reveals Drought Resistance Mechanisms of AP2/ERF Transgenic Rice

Hongryul Ahn¹, Inuk Jung², Seon-Ju Shin³, Jinwoo Park¹, Sungmin Rhee¹, Ju-Kon Kim⁴, Woosuk Jung⁵, Hawk-Bin Kwon^{3*} and Sun Kim^{1,2,6*}

¹ Department of Computer Science and Engineering, Seoul National University, Seoul, South Korea, ² Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, South Korea, ³ Department of Biomedical Sciences, Sunmoon University, Asan, South Korea, ⁴ Graduate School of International Agricultural Technology and Crop Biotechnology Institute/GreenBio Science and Technology, Seoul National University, Seoul, South Korea, ⁵ Department of Applied Bioscience, Konkuk University, Seoul, South Korea, ⁶ Bioinformatics Institute, Seoul National University, Seoul, South Korea

OPEN ACCESS

Edited by:

Zhulong Chan,
Huazhong Agricultural University,
China

Reviewed by:

Mingqiu Dai,
Huazhong Agricultural University,
China

Roel C. Rabara,
New Mexico Consortium,
United States

*Correspondence:

Sun Kim
sunkim.bioinfo@snu.ac.kr
Hawk-Bin Kwon
hbkwon@sunmoon.ac.kr

Specialty section:

This article was submitted to
Plant Abiotic Stress,
a section of the journal
Frontiers in Plant Science

Received: 27 February 2017

Accepted: 30 May 2017

Published: 15 June 2017

Citation:

Ahn H, Jung I, Shin S-J, Park J, Rhee S, Kim J-K, Jung W, Kwon H-B and Kim S (2017) Transcriptional Network Analysis Reveals Drought Resistance Mechanisms of AP2/ERF Transgenic Rice.
Front. Plant Sci. 8:1044.
doi: 10.3389/fpls.2017.01044

This study was designed to investigate at the molecular level how a transgenic version of rice “Nipponbare” obtained a drought-resistant phenotype. Using multi-omics sequencing data, we compared wild-type rice (WT) and a transgenic version (*erf71*) that had obtained a drought-resistant phenotype by overexpressing *OsERF71*, a member of the AP2/ERF transcription factor (TF) family. A comprehensive bioinformatics analysis pipeline, including TF networks and a cascade tree, was developed for the analysis of multi-omics data. The results of the analysis showed that the presence of *OsERF71* at the source of the network controlled global gene expression levels in a specific manner to make *erf71* survive longer than WT. Our analysis of the time-series transcriptome data suggests that *erf71* diverted more energy to survival-critical mechanisms related to translation, oxidative response, and DNA replication, while further suppressing energy-consuming mechanisms, such as photosynthesis. To support this hypothesis further, we measured the net photosynthesis level under physiological conditions, which confirmed the further suppression of photosynthesis in *erf71*. In summary, our work presents a comprehensive snapshot of transcriptional modification in transgenic rice and shows how this induced the plants to acquire a drought-resistant phenotype.

Keywords: rice, drought stress, drought tolerance, transcription factors, network analysis, NGS data analysis

INTRODUCTION

Plants respond to abiotic stress in various ways. By utilizing high-throughput technologies, such as microarrays and sequencing, changes in transcript levels can be detected by measuring transcripts at multiple time points as abiotic stress continues. Thus, we now have unprecedented opportunities to associate phenotypical changes with molecular-level changes in the cell. For example, the RNA sequencing (RNA-seq) technique has been used to measure expression profiles of tens of thousands of genes (37,869 gene loci in rice with evidence of expression up until 2013; Sakai et al., 2013) simultaneously under various conditions.

One important research problem that can be investigated by using the new technologies is to understand how plants respond to drought conditions. Characterizing the effects of drought-induced stress in the agricultural business has historically been a major research topic related to the productivity of crops (Venuprasad et al., 2007; Jeong et al., 2010; Ambavaram et al., 2014). Drought has become a more serious issue due to recent dramatic climate changes. The Intergovernmental Panel on Climate Change (IPCC) estimates that the global mean temperature will increase somewhere between 1 and 4°C until 2,050, and that climate-change-related drought will significantly reduce crop productivity in Africa and Asia (IPCC, 2014). Therefore, developing strategies for handling drought stress is a very important issue in plant science and also in maintaining sustainable agricultural systems.

Leveraging technological advances, plant scientists have performed analyses of time-series gene expression profiles under drought stress and reported hundreds of drought-responsive genes in rice (Rabbani et al., 2003; Rodriguez et al., 2006; Gorantla et al., 2007; Zhou et al., 2007; Rabello et al., 2008; Degenkolbe et al., 2009). So far, plant scientists have discovered many genes that play important roles in mechanisms underlying responses to drought stress, such as stomatal closure (Huang et al., 2009), osmoprotectant synthesis (Ge et al., 2008), aquaporins (Lian et al., 2004; Guo et al., 2006), cuticular wax biosynthesis (Islam et al., 2009), and phytohormone response (Iuchi et al., 2001; Qin and Zeevaart, 2002; Du et al., 2010). Additionally, a large number of drought-induced transcription factor (TF) families are reported to control drought stress responses transcriptionally (Shinozaki and Yamaguchi-Shinozaki, 2007).

However, our knowledge of biological mechanisms related to drought stress is still limited (Debnath et al., 2011; Hadiarto and Tran, 2011). In particular, we have not investigated thoroughly how drought-stress-related genes are regulated. Understanding functions, communication, and interactions of drought-stress-related genes is critical for understanding how plants respond to drought stress. The widely used differentially expressed gene (DEG) analysis, one of the single-gene analyses, is often limited in deducing meaningful biological interpretations since it does not consider relationships among genes (Shojaie and Michailidis, 2009). Fortunately, there has been significant progress in network-based analysis techniques (Barabasi and Oltvai, 2004; Barabasi et al., 2011; Gitter et al., 2013; Ma et al., 2014) that consider complex relationships among genes. These powerful network-based analysis techniques have yet to be used for analyzing omics data for rice in the context of TFs and their target genes.

In this study, we used rice plants overexpressing *OsERF71* to study drought resistance mechanisms. The *OsERF71* gene belongs to the AP2/ERF family of transcription factors. AP2/ERF is one of the well-known TF families that regulate drought-responsive genes (Chen et al., 2002; Shinozaki et al., 2003). The members share a highly conserved DNA-binding domain known as the AP2/ERF domain (Nakano et al., 2006). Members of the AP2/ERF family have been shown to exhibit diverse functions in cellular processes, such as flower development (Elliott et al., 1996), spikelet meristem determinacy (Chuck et al.,

1998), leaf epidermal cell identity (Moose and Sisco, 1996), embryo development (Boutilier et al., 2002), and stress tolerance (Dubouzet et al., 2003). Moreover, overexpression of AP2/ERF transcription factors enhances tolerance to drought in several plant species: *Arabidopsis*, tomato, wheat, and rice (Haake et al., 2002; Hsieh et al., 2002; Pellegrineschi et al., 2004; Oh et al., 2005, 2009). Recently JK Kim and his colleagues (Lee et al., 2016) reported that rice plants overexpressing the *OsERF71* TF showed enhanced survival over the wild-type under drought stress at the vegetative stage of growth and a 23–42% increase in total weight gain over the wild-type under drought stress at the reproductive stage of growth. They also investigated the target genes of *OsERF71* and found that *OsERF71* regulates *OsCINNAMOYL-COENZYME A REDUCTASE1*, a gene playing a role in lignin biosynthesis, and that rice plants overexpressing *OsERF71* formed enlarged aerenchyma and had high lignification levels.

To study drought resistance mechanisms, we investigated transcriptomic changes in response to dehydration stress between wild-type rice (WT) and *OsERF71*-overexpressing rice (*erf71*) lines. The goal was to elucidate the association between the *erf71*-specific response at the transcriptome level and the drought-resistant phenotype. For the investigation, we generated comprehensive multi-omics data, such as RNA-seq data, micro RNA (miRNA) sequencing data, and whole-genome bisulfite DNA methylation sequencing data, collected at 0, 1, and 6 h after treatment (HAT) under dehydration stress (i.e., aeration and without watering), from the two rice genotypes. A comprehensive bioinformatics analysis pipeline, including TF networks and a cascade tree, was developed and used for the analysis of the multi-omics data.

MATERIALS AND METHODS

Plant Material and Drought-Stress Treatments

We germinated seeds of WT (*Oryza sativa* ssp. *japonica* “Nipponbare”) plants and rice plants overexpressing *OsERF71* (Os06g0194000) in a “Nipponbare” background (Lee et al., 2016) on petri-dishes and then cultured them in Yoshida’s solution (Yoshida et al., 1976) maintained in a temperature-controlled culture room at 29°C under 16/8-h light/dark conditions. Rice plants at the three-leaf stage were subjected to dehydration stress by removal of the culture solution. Untreated plants were used as a control. After treatment, entire plants were immediately transferred into liquid nitrogen. After a pilot experiment to measure the expression of the *Dip1* gene in WT using reverse-transcription (RT)-PCR, we selected 0, 1, and 6 HAT as time points to measure omics data. The whole plants harvested at these time points were kept frozen in liquid nitrogen until DNA/RNA extraction.

Measuring Gene Expression Levels

Our eight mRNA-seq data sets (0, 0.5, 1, 3, and 6 HAT mRNA-seq data for WT, and 0, 1, 6 and HAT mRNA-seq data for *erf71* upon drought stress) were processed accordingly. After total RNA was extracted using Tri-Reagent (MRC, Cincinnati, OH,

USA), poly(A) RNAs were isolated using poly-T oligo-attached magnetic beads. These were next fragmented, and primer was attached to them. cDNA strands were then synthesized, and an adaptor was ligated. After PCR amplification, an mRNA library was prepared. mRNA deep sequencing (TruSeq RNA sample preparation Guide, Illumina) produced short reads. After clipping adapter/primer sequences from the short reads, we mapped them to the IRGSP1.0 reference genome (Kawahara et al., 2013; Sakai et al., 2013) using Tophat (Trapnell et al., 2009). We quantified the expression level of each gene using Cufflinks (Trapnell et al., 2010) based on gene information that was downloaded from the Rice Annotation Project website (<http://rapdb.dna.affrc.go.jp/>). DEG analysis was then performed using cuffdiff in the Cufflinks package.

Network-Based Characterization of *OsERF71* Transgenic Rice

To investigate how overexpression of the *OsERF71* TF gene affected other genes, we used a TF-network-based analysis. The challenge here was that we had limited knowledge of the relationships between the TF and its target genes. In addition, a TF activates or suppresses other TFs, producing chains of activation and suppression events. Hence, we developed a two-step bioinformatics strategy of TF-network-based analysis, as described in **Figure 1**.

- **Step 1:** Constructing a dehydration TF network utilizing gene expression data from databases and a dehydration experiment.
- **Step 2:** Instantiating phenotype-differential dehydration networks and identifying DEG modules.

Network Analysis Step 1: Constructing a Dehydration TF Network Utilizing Gene Expression Data from Databases and a Dehydration Experiment

To construct a reliable dehydration TF network, we first constructed a template TF network utilizing large-scale microarray data. Before constructing a template TF network, factors, such as choice of network construction method, the data set size, and cutoff values were thoroughly investigated. A recent study investigated many network construction methods and reported that mutual-information and correlation-based methods recovered feed-forward loops most reliably (Marbach et al., 2012). Since the goal of this study was to investigate the effect of *OsERF71*-overexpression on other genes through relationships between the TF and its target genes, which can be seen as a feed-forward propagation from *OsERF71* to other genes, we selected Pearson's correlation coefficient (PCC) as the network construction method. Large-scale gene expression data sets (1,893 microarray data sets) were downloaded from the OryzaExpress Gene Expression Network website (<http://bioinf.mind.meiji.ac.jp/OryzaExpress/>). Probe-IDs that were used for microarray experiments were converted to gene-IDs according to a previous study (Miller et al., 2011). Since PCC is shown to converge as the sample size increases, we performed an empirical study to determine whether the data set size

was beyond the convergence threshold and was sufficient to construct a template TF network robustly. Varying the number of samples, we produced different sample-size subsets of the microarray data by random sampling from the 1,893 microarray data sets. For each subset, PCCs between TF and target genes were then computed and PCC density distributions and network topologies were investigated (see "TF network construction" in Discussion). We observed that the density distributions and the network topologies converged with a sample size greater than approximately 800. This showed that 1,893 microarray data sets were sufficient to produce a robust template TF network. Recent studies that used the PCC method for biological network construction detected modular structures of genes in *Arabidopsis*, rice, and maize networks (Mao et al., 2009; Ficklin and Feltus, 2011). These studies reported that each of the modules had a specific biological function. Based on this result, we defined functionality score as follows.

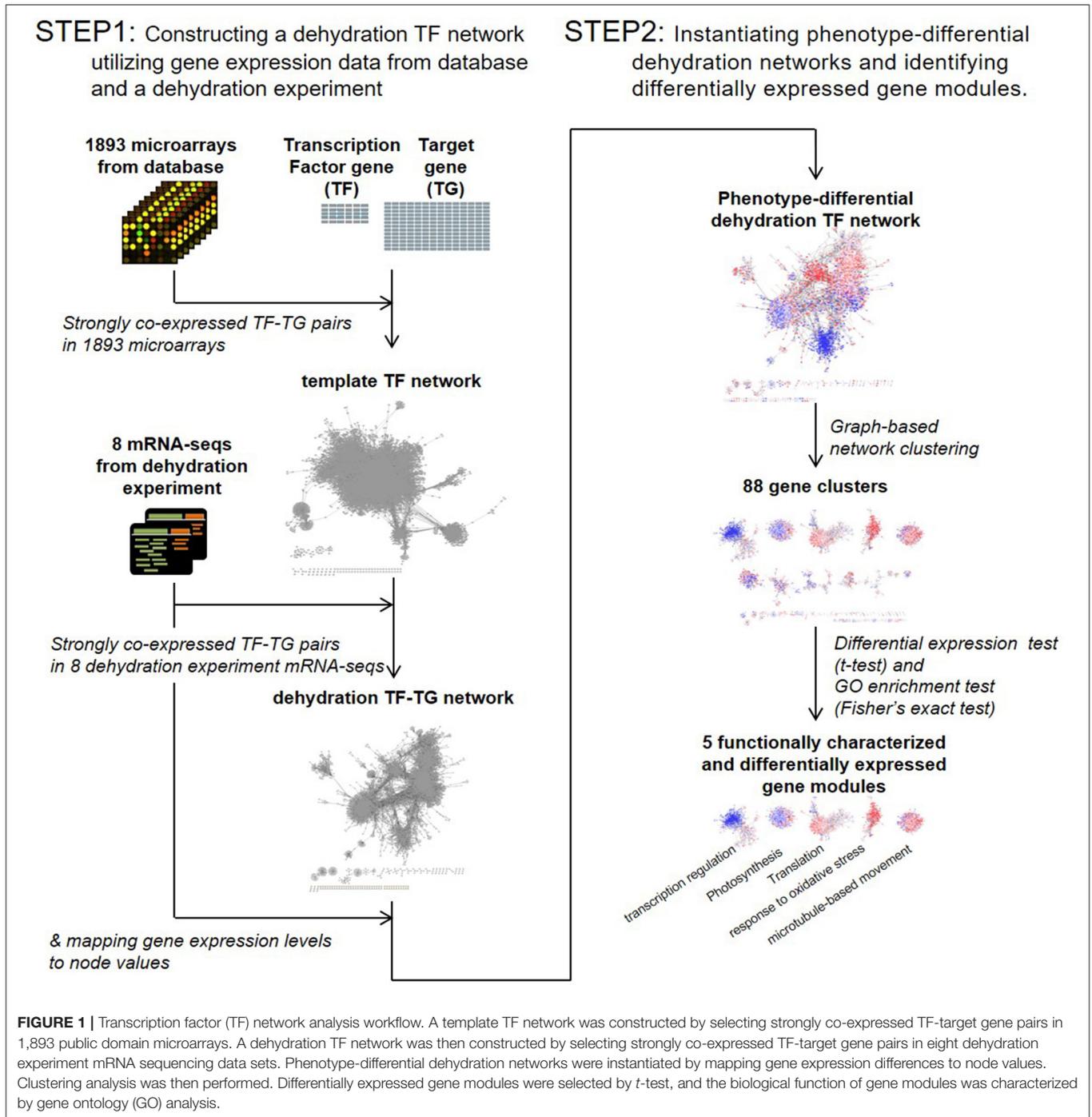
$$\text{Functionality Score}(G) = - \sum_{c_i \in C} \frac{|c_i|}{N} \log_{10}(p_{c_i})$$

In the formula, G is a network and it is divided into a set of gene clusters, $C = [c_1, c_2, \dots, c_n]$, using a graph-based clustering algorithm (Blondel et al., 2008). N is the number of genes in the network, and p_{c_i} is the p -value of the most significant gene ontology (GO) term in a GO enrichment test of the cluster c_i . The functionality score measures how well a network is divided into functional gene modules. We investigated functionality score for each network constructed at different PCC cutoff values (see "TF network construction" in Discussion). The cutoff value of 0.67 was chosen because the functionality score was maximized at the cutoff value. TF-target gene pairs with strong association ($|PCC| > 0.67$) in the 1,893 microarray data sets were then defined as edges in the template TF network. The template TF network consisted of 10,740 genes (898 TFs and 9,842 nonTFs) and 135,550 links (4,073 TF-TF links and 131,477 TF-nonTF links). A dehydration TF network was then constructed by selecting edges in the template TF network that had strong association ($|PCC| > 0.67$) in eight dehydration experiment mRNA-seq data sets. The constructed dehydration TF network consisted of 7,319 genes (729 TFs and 6,590 nonTFs) and 50,672 links (1,375 TF-TF links and 49,297 TF-nonTF links). The topology of the dehydration TF network is shown in Supplemental Data Set 1. The topology of the network was visualized using Cytoscape (Saito et al., 2012).

Network Analysis Step 2: Instantiating Phenotype-Differential Dehydration Networks and Identifying Differentially Expressed Gene Modules

In this step, our goal was to identify DEG modules between WT and *erf71* from the dehydration TF networks. To do so, we employed the following strategy.

- **Step 2-1:** Phenotype-differential dehydration TF networks were instantiated by mapping gene expression differences to node values of the dehydration TF network.



- **Step 2-2:** Graph-based network clustering broke down the phenotype-differential dehydration TF networks into several gene clusters according to connectivity.
- **Step 2-3:** Differential expression and GO enrichment tests were performed for each cluster.
- **Step 2-4:** DEG modules were selected and designated as “modules.”

In Step 2-1, we instantiated phenotype-differential dehydration TF networks by mapping gene expression differences

between time points for each plant (i.e., $\log_2(W1/W0)$ and $\log_2(W6/W0)$) as well as differences across plants (i.e., $\log_2(E1/A0) - \log_2(E1/W0)$) to nodes of the dehydration TF network, where “W” and “E” stand for WT and *erf71*, and “0”, “1,” and “6” stand for 0, 1, and 6 HAT, respectively. In Step 2-2, the phenotype-differential dehydration TF networks were broken down into several gene clusters using a multi-level network clustering method (Blondel et al., 2008) that groups highly connected nodes into a cluster of nodes. In Step

2–3, a paired sample *t*-test was performed on each cluster to determine whether each gene cluster was differentially expressed or not. Also, biological functions of each cluster were characterized by GO enrichment analysis based on Fisher's exact test. In Step 2–4, the clusters showing high-level significance ($p < e^{-9}$) in both tests were selected and designated as "modules."

Measuring Photosynthesis Levels Using a Liquid CO₂ Supply System to Confirm Suppression of Photosynthesis under Drought Stress

Since the suppression of photosynthesis was a key observation from analyzing sequencing data, we confirmed photosynthesis levels using a liquid CO₂ supply system. For this confirmation experiment, we re-grew rice while measuring CO₂ levels. Since we needed a large amount of CO₂, we had to change experimental conditions from the three-leaf stage to the eight-leaf stage of rice and also the time points used. We selected time points corresponding to the same physiological status and degree of dehydration stress by measuring gene expression levels of *Dip1* through RT-PCR.

For photosynthesis measurements, rice plants at the eight-leaf stage were subjected to drought stress for 24, 48, and 72 h by removal of the culture solution. For measurements of Pnet and Cs, a Li-Cor 6400 photosynthesis measurement system (Li-Cor, Lincoln, NE) was used to measure gas-exchange ability. Gas-exchange measurements were performed at four stages: before treatment, and 24, 48, and 72 h after drought stress treatment. Leaves were measured in 2 × 3 cm photosynthesis measuring cuvettes under 1,500 mol m⁻² s⁻¹ light emitting diode (LED) light, 400 mmol L⁻¹ CO₂ and a cuvette temperature of 25°C. Measurements were conducted using the youngest among the 3–4 fully expanded leaves of the WT and *erf71* transgenic plants. Once a leaf was clamped in the chamber, data were automatically collected every 10 s for 3 min, while mid-point measurements were collected as representative data. A liquid CO₂ supply system (Li-Cor 6400-01, Lincoln, NE) attached to the photosynthesis measurement system was used to measure photosynthetic changes based on changes in supplied CO₂ concentration. CO₂ concentration was adjusted to 400, 300, 200, 100, 50, and 0 mmol L⁻¹ CO₂. The conditions in the cuvette were 1500 mol m⁻² s⁻¹ photosynthetically active radiation, 25°C, and 30–60% relative humidity.

Constructing a TF Cascade Tree with *OsERF71* as the Root Node from the Dehydration TF Networks

Dehydration differential TF networks were converted into an *OsERF71*-centric cascade tree to investigate the connection between the *OsERF71* transgene and the five modules that were identified in the TF network analysis. To achieve this goal, we added new edges of strong association ($|PCC| > 0.5$) to the TF network to compensate for the loss of connectivity between the *OsERF71* transgene and the five modules. The strong association value of 0.5 was chosen because it was the lower boundary

that was used for network construction in previous rice species network studies (Lim et al., 2010; Zhang L. et al., 2012; Lu et al., 2013; Yang et al., 2013; Hwang et al., 2014). We confirmed that the five modules were maintained after addition of the new edges by measuring the clustering coherency value (Fowlkes and Mallows, 1983). The network was then converted into an *OsERF71*-centric cascade tree through level-wise construction of relationships starting at *OsERF71*. In this process, *OsERF71* was located to the top of the tree and the direct target genes of *OsERF71* in the TF network were then connected with *OsERF71* in the cascade tree and located at the level below in the tree. This process was performed repeatedly for all genes of the five modules. The full landscape of interactions in the constructed tree was very complex and it included multiple paths to a single gene and back edges (i.e., the edges from the TF to the target gene of previous levels). This complex relationship needed to be simplified to focus the effect of *OsERF71* overexpression. Therefore, we selected the single path with the greatest PCC score and removed the other paths as well as the back edges. The topology of the TF cascade tree is shown in Supplemental Data Set 2.

Measuring miRNA Expression Levels and Inferring miRNA-Gene Interaction

We prepared a small RNA library by attaching 3' and 5' adaptors to total RNA, amplifying them by PCR and performing gel purification. After deep sequencing (Illumina's TruSeq™ small RNA sample prep kit), 51-nt-length reads were generated. After the adaptor sequence was removed, miRanalyzer (Hackenberg et al., 2011) was used to quantify the expression of known rice miRNA sequences from the miRBase database (Kozomara and Griffiths-Jones, 2014). The number of reads matching known miRNAs divided by the total aligned reads was used to determine the expression level of miRNAs. Using psRNATarget, a webserver for predicting small RNA targets in plants (Dai and Zhao, 2011), candidate miRNA-target gene pairs were generated. For each of those candidates, the miRNA-target gene pairs showing a strong negative association of expression ($PCC < -0.67$) in our dehydration experiment were selected for the final miRNA-target gene interactions.

Measuring DNA Methylation Levels and Identifying Differentially Methylated Regions

After isolating DNA, we performed whole-genome bisulfite sequencing from samples of WT and *erf71* collected at 0, 1, and 6 HAT. Adaptor sequences were removed from the raw reads, and the sequences were aligned to the rice genome using BS-MAP (Xi and Li, 2009). CpG sites with two reads or fewer aligned were not considered for further analysis. For each CpG site, we determined the methylation level as the number of methyl cytosines (i.e., cytosine transformed into thymine by bisulfite treatment) aligned to the site divided by the number of reads aligned to the site. This resulted in a normalized methylation level from 0 (hypo-methylated) to 1 (hyper-methylated).

We then considered a CpG site as a differentially methylated cytosine if the difference in maximum and minimum methylation levels among six samples was greater than 0.7, a measure for differential methylation in the CpG context in previous studies (Hsieh et al., 2009; Chodavarapu et al., 2012; Stroud et al., 2013), at the CpG site. The number of differentially methylated cytosines was 163,982 (1.4%) among 11,512,277 CpG sites with at least three reads mapped. We defined a differentially methylated region (DMR) as a genomic region containing more than eight differentially methylated cytosines in succession ($p < 1.0 \times 10^{-12}$ by Poisson probability). The number of DMRs was 1,607 and they covered 0.14% (526,064 nt/373,245,519 nt) of the rice genome.

RESULTS

Time-Series Experiments and Multi-Omics Sequencing Data

We carried out a pilot study to select proper time points for measuring omics data. Yi and colleagues (Yi et al., 2010) reported that *Dip1* (Os02g0669100) is induced upon dehydration stress treatment, reaching a peak at 2 HAT and remaining constant until 8 HAT, which could be used as a response marker for water deficiency. Thus, we measured the expression level of *Dip1* using RT-PCR at different time points during the dehydration treatment. The expression level of *Dip1* peaked at 3 HAT and continued to be expressed until 6 HAT. This was also confirmed by the mRNA-seq experiment (Supplemental Figure 1). Thus, three time points, 0, 1, and 6 HAT, before and after the peak expression level of *Dip1*, were selected to study how the plants responded to dehydration stress. We first performed mRNA sequencing for six samples, i.e., WT and *erf71* for 0, 1, and 6 HAT. We also performed small RNA sequencing and whole-genome bisulfite sequencing to see if small RNAs or methylation have any major regulatory roles under drought stress in addition to TFs. In summary, we generated three types of omics datasets (mRNA-seq, small RNA sequencing, and whole-genome bisulfite sequencing) from the two rice plants at three time points under dehydration stress.

Gene Expression Profile Comparison between WT and *OsERF71*-Overexpressing Rice

A total of 18,567 and 18,555 genes with FPKM (Fragments Per Kilobase of exon per Million fragments mapped) greater than 1.0 were present in WT and *erf71* samples, respectively. The overall expression level of genes decreased as the dehydration stress continued in both WT and *erf71* (Figure 2A). The p -values were 2×10^{-45} , 1×10^{-434} , 2×10^{-80} , and 1×10^{-348} for the paired t -tests with H1: $W_0 > W_1$, $W_1 > W_6$, $E_0 > E_1$, and $E_1 > E_6$, respectively, where “W” and “E” stand for WT and *erf71*, and “0”, “1”, and “6” stand for 0, 1, and 6 HAT. This observation indicated that both WT and *erf71* responded to dehydration by lowering gene expression levels globally. DEG analysis determined DEGs as genes being differentially expressed with $p < 0.05$ in statistical tests using Cuffdiff (Trapnell et al., 2010) and showed that the number of

DEGs increased more than 3-fold between the 0-to-1 HAT period and the 0-to-6 HAT period in both plants (Figure 2B). These two results showed that the transcriptomes of both plants were affected gradually as dehydration stress continued.

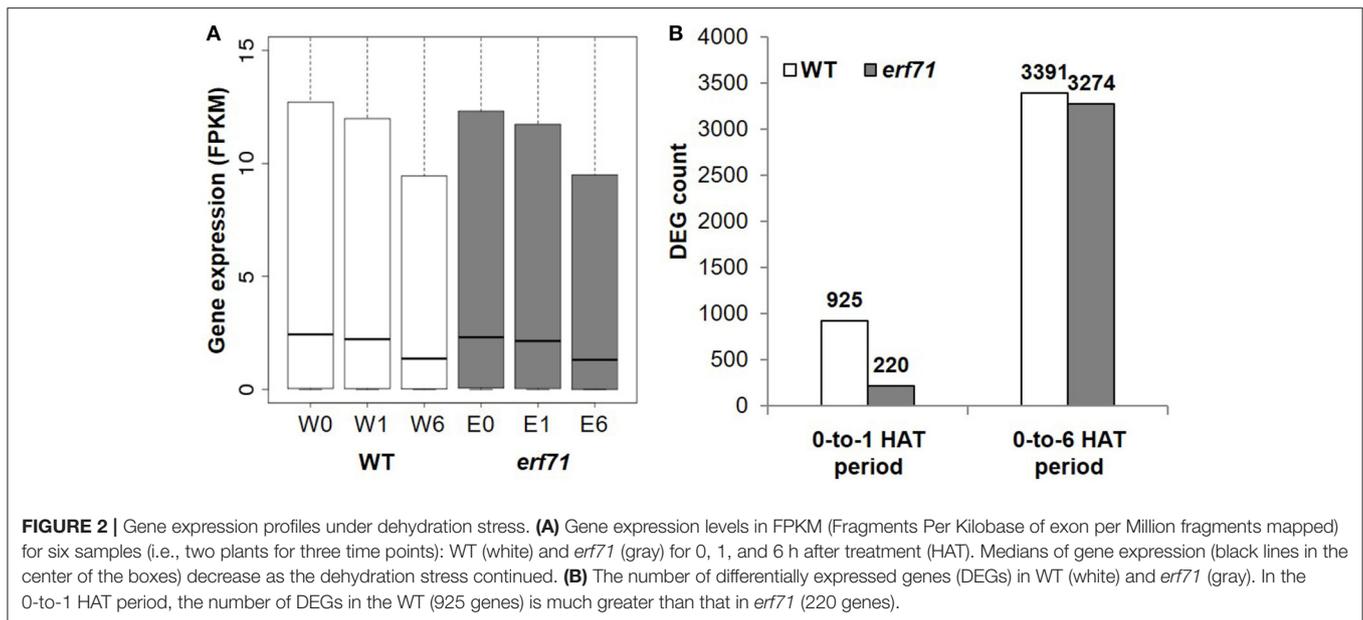
However, *erf71* showed a lesser degree of change in gene expression under dehydration stress in the 0-to-1 HAT period. The decrease in global gene expression in *erf71* was smaller than that in WT with a significance of $p < 0.05$ by paired t -test with H1: $\log(W_1/W_0) < \log(E_1/E_0)$. Most prominently, the difference in the number of DEGs between WT and *erf71* was more distinctive in the early response phase (i.e., the 0-to-1 HAT period) than in the late response phase (i.e., the 0-to-6 HAT period). The number of DEGs in *erf71* (220 genes) was four times smaller than that in WT (925 genes) in the 0-to-1 HAT period. These results suggest that *erf71* was less sensitive to dehydration stress at the global transcriptome level in the early response phase (i.e., the 0-to-1 HAT period).

We next performed a GO enrichment analysis of DEGs in the 0-to-1 HAT period by employing Fisher's exact test. We identified 16 and 10 GO terms being enriched with $p < 0.05$ for WT and *erf71*, respectively. Five GO terms were commonly found in WT and *erf71*, such as “oxidation-reduction process” and “protein ubiquitination”. Eleven GO terms were specific to WT, whereas five were specific to *erf71* (Supplemental Table 1). However, the GO terms that were specifically enriched to WT or *erf71* were not intuitive for elucidating the difference of response between WT and *erf71*, which showed the limitations of the DEG-based analysis. Recent studies reported that network analysis has benefits for analyzing genomics data at the level of individual genes (i.e., DEG analysis) because it considers relationships between genes (Chi et al., 2014) and results are often easier to interpret and potential causal mechanisms can be identified (The Mutation and Pathway Analysis Working Group of the International Cancer Genome, 2015). Thus, we developed and performed a systemic bioinformatics pipeline that utilizes a transcriptional network to characterize the difference in gene expression response between WT and *erf71*. Especially, we focused on gene regulation mechanisms by TFs since our goal was to investigate the transcriptomic influence affected by the overexpression of *OsERF71*, a TF gene.

Network-Based Characterization of *OsERF71*-Overexpressing Rice

The goal of this step was to characterize the differences in gene expression response between WT and *erf71* utilizing a TF-target gene regulatory network (TF network hereafter). In a TF network, nodes are genes (TFs and nonTFs) and edges are connections between TFs and potential target genes including TFs. Our TF network did not include edges between nonTF genes since we focused on transcription-factor-centered regulation. To achieve this goal, we developed and implemented a bioinformatics pipeline for TF network analysis (see “Network-based characterization of *OsERF71* transgenic rice” in Materials and methods).

A template TF network was constructed by selecting TF-target gene pairs with strong associations ($|PCCs| > 0.67$) in



eight mRNA-seq data sets from our experiment and in 1,893 microarray data sets in the public domain. Differences in gene expression levels (i.e., \log_2 fold change) were then assigned to the genes in the network. **Figures 3A,B** shows the constructed dehydration TF networks of WT and *erf71*, respectively, where red/blue dots denote up/down-regulated genes before and after dehydration stress (i.e., 0 HAT vs. 1 HAT). **Figure 3C** shows a phenotype-differential TF network where red/blue dots denote relatively differentially regulated genes between WT and *erf71*. In other words, red dots indicate genes that are relatively up-regulated (more up-regulated or less down-regulated) in *erf71* under dehydration stress.

We divided the TF network into 88 gene clusters by grouping highly connected genes into a cluster by a graph-based network clustering algorithm. We then characterized each cluster by differential gene expression and GO enrichment tests. Finally, we identified five gene clusters, with 713, 1,363, 537, 1,586, and 1,605 genes, showing high-level significance ($p < e^{-9}$) in both tests. We designated the five clusters as “modules.” **Figure 4** shows the characteristics of the five modules—plots of the expression levels, the position in the TF-TG network, and the results of GO enrichment tests.

As shown in **Figure 4A**, the mean expression levels of the five modules were changed in one direction (i.e., increased or decreased) as dehydration stress continued in both rice plants, but the degree of change was different between the two rice plants. Module 1 was up-regulated in both types of rice but less up-regulated in *erf71*. Modules 2, 3, and 4 were down-regulated in both types of rice but less down-regulated in *erf71*. Module 5 was down-regulated in both types of rice but more down-regulated in *erf71*.

The results of GO enrichment analysis showed that each module had different enriched GO biological process terms. Module 1 included drought-response-related TFs. A majority of genes in Module 2 were related to translation; most of the genes

in Module 3 were related to response to oxidative stress; Module 4 included genes with diverse functions but many genes were related to the cell division cycle; and a large number of genes in Module 5 were related to photosynthesis. **Figure 4C** summarizes the results of the differential gene expression test and GO terms enrichment test for the five modules. The next section presents detailed analysis on each module.

Analysis of Module 1: Drought-Response-Related TFs Are Up-Regulated Less in *OsERF71*-Overexpressing Rice

The GO term highly enriched in Module 1 was DNA-dependent regulation of transcription (GO:0006355). According to the TF list obtained from the plant TF special database, PlantTFDB (Jin et al., 2014), about a fifth of the genes in Module 1 (143/713) consisted of TFs: WRKY (27), ERF (23), NAC (17), C2H2 (10), bZIP (9), bHLH (7), MYB (6), GRAS (6), HSF (5), MYB related (5), Trihelix (4), C3H (4), HD-ZIP (2), Dof (2), NF-YB (2), SBP (2), G2-like (2), ARR-B (2), NF-YC (1), CO-like (1), RAV (1), CAMTA (1), VOZ (1), ARF (1), CPP (1), and DBB (1), where the numbers in parentheses indicate the number of genes included in the module. Among them, TF families, such as WRKY, ERF, NAC, C2H2, bZIP, bHLH, and MYB are well-known drought-stress-related TF families (Chen et al., 2012; Mizoi et al., 2012; Nakashima et al., 2012, 2014; Lindemose et al., 2013; Singh and Laxmi, 2015). Moreover, alterations in the expression of 10 TFs in Module 1, Os01g0797600 (*OsAP37*) (Oh et al., 2009), Os01g0968800 (*OsDREB1F*) (Wang et al., 2008), Os02g0654700 (*OsAP59*) (Oh et al., 2009), Os03g0741100 (*OsbHLH148*) (Seo et al., 2011), Os03g0815100 (*SNAC1*) (Hu et al., 2006), Os03g0820300 (*ZFP182*) (Zhang H. et al., 2012), Os05g0322900 (*OsWRKY45*) (Qiu and Yu, 2009), Os11g0127600 (*ONAC045*) (Zheng et al., 2009), Os11g0184900 (*OsNAC5*)

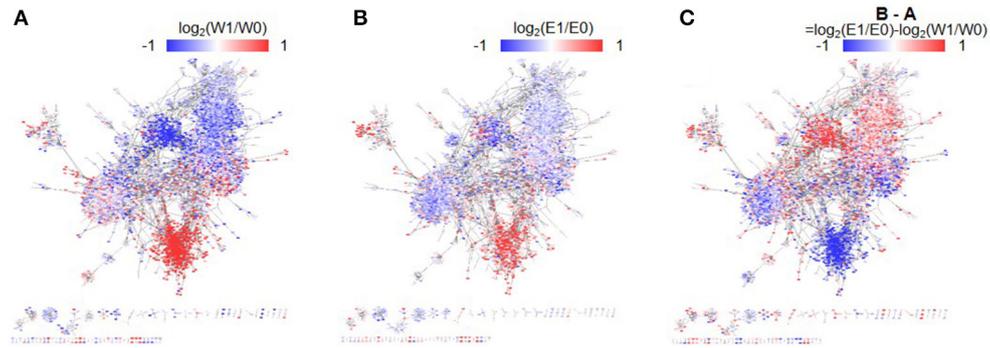


FIGURE 3 | Phenotype-differential dehydration transcription factor (TF) networks. The three dehydration differential networks were instantiated by mapping gene expression differences to node values of the dehydration TF network. The node values are represented by red-white-blue-gradation. The two time-point differential networks (**A,B**) were instantiated by mapping gene expression differences between time points, such as $\log_2(W1/W0)$ and $\log_2(E1/E0)$, respectively, where “W” and “E” stand for WT and *erf71*, and “0,” “1,” and “6” stand for 0, 1, and 6 h after treatment. In these networks, the red/blue color represents up-/down-regulation of gene expression under dehydration stress. A phenotype-differential network (**C**) was instantiated by mapping gene expression differences between the two rice plants, such as $\log_2(E1/E0) - \log_2(W1/W0)$. In this network, the red/blue color represents relative up-/down-regulation of gene expression in *erf71* compared with WT. The dehydration differential TF networks showed gene cluster structures distinctively, where a gene cluster indicates a sub-network with member genes highly connected to each other. They also showed a trend that member genes within gene clusters had common gene expression difference patterns.

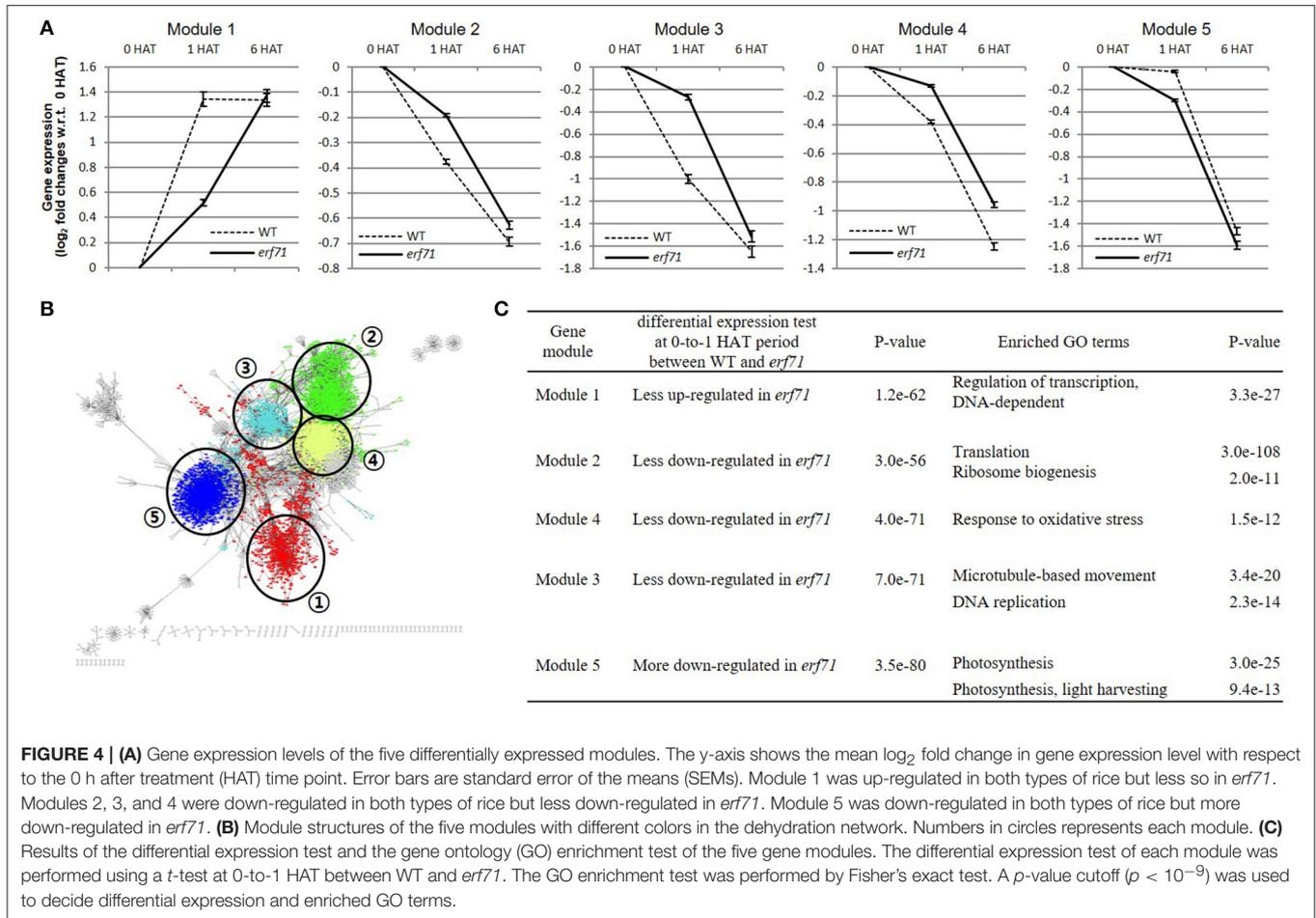


FIGURE 4 | (A) Gene expression levels of the five differentially expressed modules. The y-axis shows the mean \log_2 fold change in gene expression level with respect to the 0 h after treatment (HAT) time point. Error bars are standard error of the means (SEMs). Module 1 was up-regulated in both types of rice but less so in *erf71*. Modules 2, 3, and 4 were down-regulated in both types of rice but less down-regulated in *erf71*. Module 5 was down-regulated in both types of rice but more down-regulated in *erf71*. **(B)** Module structures of the five modules with different colors in the dehydration network. Numbers in circles represents each module. **(C)** Results of the differential expression test and the gene ontology (GO) enrichment test of the five gene modules. The differential expression test of each module was performed using a *t*-test at 0-to-1 HAT between WT and *erf71*. The GO enrichment test was performed by Fisher’s exact test. A *p*-value cutoff ($p < 10^{-9}$) was used to decide differential expression and enriched GO terms.

(Song et al., 2011), and Os12g0583700 (*ZFP252*) (Xu et al., 2008), have already been reported to produce drought-resistant phenotypes.

We conducted RT-PCR experiments to confirm the expression levels of the TFs in Module 1 that have not yet been documented to affect drought resistance.

Among them, eight TF genes, Os02g0764700 (*OsERF103*), Os03g0180900 (*TIFY11C*, *OsJAZ2*), Os03g0327100 (*ONAC039*, *OsCUC1*), Os03g0820400 (*ZFP15*), Os04g0671800 (*OsC3H32*), Os04g0676700, Os06g0670300, and Os12g0123800 (*ONAC132*, *ONAC300*), were shown to be up-regulated in both plants but less up-regulated in *erf71* during the 0-to-1 HAT period in response to drought stress (Supplemental Figure 2). Os03g0180900 (*TIFY11C*, *OsJAZ2*) is a gene of the JAZ family that contains a well-conserved domain called ZIM or TIFY (Vanholme et al., 2007). It was induced under drought stress and also by the overexpression of *OsbHLH148*, a gene that causes drought tolerance when overexpressed. Also, *OsJAZ2* exhibited a weak interaction with *OsbHLH148* and has been proposed to target activation of *OsbHLH148* (Seo et al., 2011). Os03g0327100 (*ONAC039*, *OsCUC1*) and Os12g0123800 (*ONAC132*, *ONAC300*) were reported to be responsive to drought, salt, and cold stress (Fang et al., 2008). These results show that the eight TFs were differentially regulated in WT and *erf71* during the 0-to-1 HAT period and putatively related to the drought-resistance mechanism.

Genes in Module 1 were up-regulated in both WT and *erf71*, and Module 1 was the only up-regulated module among all five modules. As overall gene expression levels decreased with continued dehydration stress, indicating the suppression of various activities, up-regulation was a relatively unexpected phenomenon. We found that Module 1 contained many up-regulated DEGs (31 up-DEGs among total 112 up-DEGs in both plants in the 0-to-1 HAT period) with a significance level of $p < 1.0e-23$ by Fisher's exact test. In summary, the expression level of TF genes was increased in our dehydration response network, and those TF genes seemed to form a modular structure in the network clustering analysis, suggesting that there might be particular biological functions of the TF module. This observation needs further investigation, which was beyond the scope of this study.

OsERF71, the overexpressed gene, was present in Module 1 and it had three direct neighbors (Os03g0701700, Os10g0346600, and Os11g0157200) in the module. Genes in Module 1 were directly connected to the transgene, unlike those in the other modules.

Although gene expression levels increased in Module 1, the degree of change differed between WT and *erf71*. Gene expression increased less in *erf71* in the early response phase (i.e., the 0-to-1 HAT period). This trend, relatively small gene expression changes in *erf71* in the early response phase, was observed consistently in the results of other analyses: the number of DEGs was smaller in *erf71*, and the magnitude of change in expression was smaller for genes in Modules 2, 3, and 4 and globally during the 0-to-1 HAT period.

Analysis of Modules 2, 3, and 4: Critical Survival-Related Genes Are Maintained

Three modules, Module 2, Module 3, and Module 4, included genes that were relatively up-regulated in *erf71* compared with WT. The enriched GO terms were genetic information processing and translation for Module 2; response to oxidative

stress for Module 3; and cell cycle for Module 4. All significantly enriched GO terms were commonly related to essential biological processes for sustaining life.

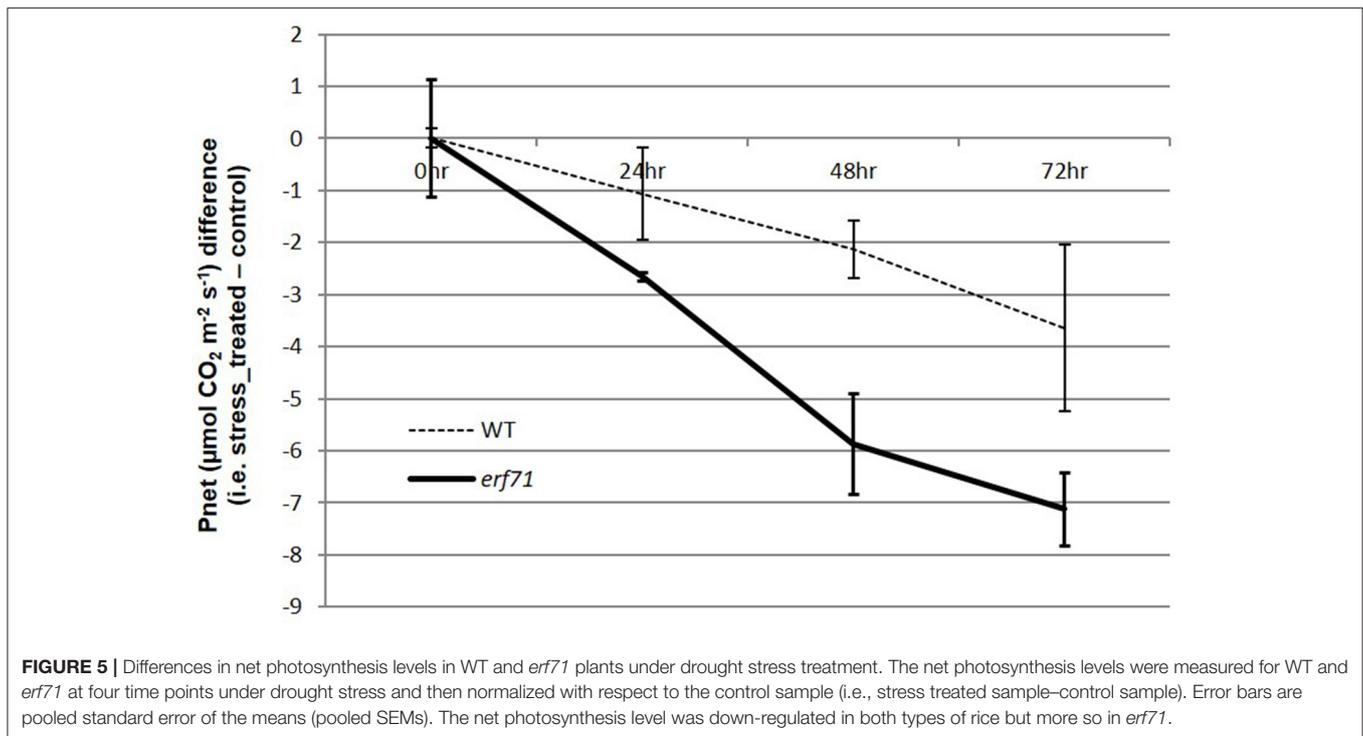
In Module 3, 54 genes were related to oxidative reduction (GO:0055114) while 23 were related to response to oxidative stress (GO:0006979). Oxidation is closely related to water deficiency tolerance in plants. In particular, reactive oxygen species (ROS) are known to be overproduced in response to abiotic stress. ROS are highly reactive and toxic, causing damage to proteins, lipids, carbohydrates, and DNA when they exceed the cell's antioxidant removal capacity (Miyamoto et al., 2003; Gill and Tuteja, 2010). Since those genes in *erf71* were down-regulated to a lesser extent than in WT, it is possible that *erf71* is more capable of detoxifying the rising level of oxidation, preventing severe damage to the plant.

Analysis of Module 5: Expression of Genes That Are Related to Photosynthesis Is Down-Regulated Further

Module 5 consisted of genes that were down-regulated more in *erf71* compared with WT. The significant GO terms enriched in the module were related to photosynthesis ($p < e-13$). During photosynthesis, the plant synthesizes chemical compounds using energy from light. However, such photosynthetic metabolic processes require energy use by the plant. For example, toxic elements are generated as a subsidiary product that must be detoxified, requiring the production of anti-toxic elements by the plant. Thus, maintaining such photosynthetic metabolism during a critical situation, such as dehydration, will hinder the survival of the plant (Ramachandra Reddy et al., 2004). In our analysis, *erf71* transgenic rice showed strong down-regulation of gene expression levels of photosynthetic genes compared with WT, suggesting that *erf71* was possibly able to shut down photosynthesis mechanisms in response to dehydration stress.

Photosynthesis Is Suppressed Physiologically Further in *OsERF71*-Overexpressing Rice

Smaller changes in gene expression in *erf71* in the early response phase (i.e., the 0-to-1 HAT period) were observed consistently. For instance, the number of DEGs was smaller in *erf71*, and the magnitude of the decrease in expression was smaller in *erf71* when considering all genes. TF network analysis also showed that genes in Modules 2, 3, and 4 were related to survival-associated biological functions under stress conditions, such as microtubule-based movement, translation, and response to oxidative stress, and these were down-regulated less in *erf71* compared with WT. This observation is intuitive since maintaining gene expression levels of survival-related genes promotes the dehydration-resistant phenotype. However, genes in Module 5 that were related to photosynthesis showed a greater response in *erf71* (i.e., the genes in Module 5 were down-regulated more in *erf71*). Since this was a key observation in this study, we measured the photosynthetic levels of WT and *erf71* plants under dehydration stress through an experiment at the physiological level. The experiment confirmed that net



photosynthesis levels decreased in both plants but with greater magnitude (2-fold) in *erf71* (Figure 5).

OsERF71 Cascade Tree Analysis at 1 and 6 HAT

We investigated the effects of overexpression of *OsERF71* in *erf71* by converting the differential TF regulatory network of five gene modules described in the previous section into an *OsERF71*-transgene cascade tree. The *OsERF71*-transgene cascade tree is a tree graph structure with *OsERF71* as the root node and putative downstream genes in the five modules as child nodes. The network was transformed into a tree by applying the single-source shortest TF regulatory path strategy (see “Constructing TF cascade tree with *OsERF71* as the root node from dehydration differential TF networks” in Materials and methods). The resulting *OsERF71* cascade tree is shown in Figure 6. We defined “depth” in the tree as the number of edges in the shortest path from *OsERF71* to the gene. Small-world networks are known to have a property called six degrees of separation, in which all nodes in the network are connected to each other within a few steps. The cascade tree preserves this property since *OsERF71* is connected to all the genes within six depth levels in our cascade tree. In addition, we found in the cascade tree, major paths starting at the *OsERF71* transgene followed by regulatory hub TFs. Here, regulatory hub TFs (described by the large rectangle nodes in Figure 6) indicated transcription factors that had a large number of genes (more than 200 genes) in their downstream in the cascade tree. Moreover, we observed that there were nine dominant major paths covering the majority of genes of each module in the cascade tree. For

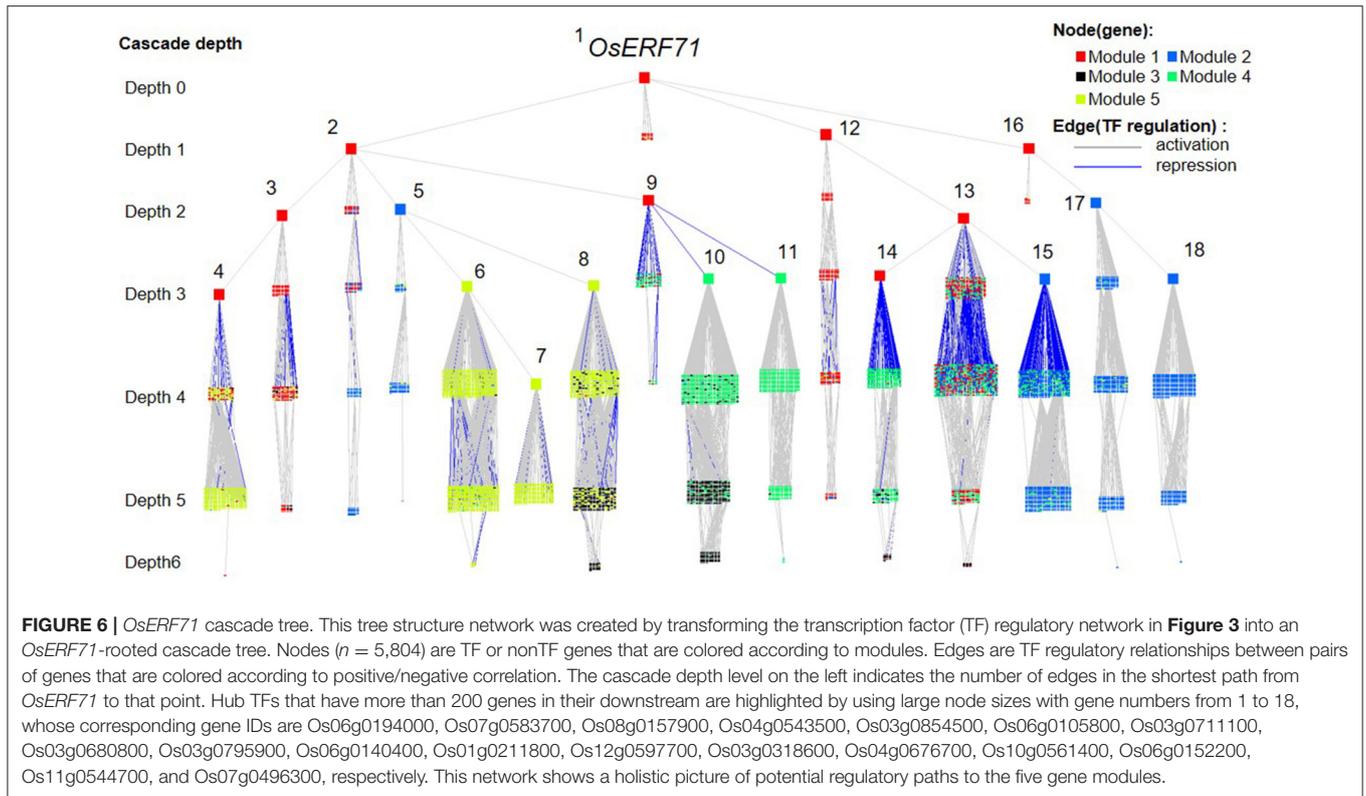
instance, 79% of the genes in Module 5 were downstream of two major regulatory paths. One of the paths included 56% of the genes of Module 5. The upstream of the path consisted of a chain of hub TFs, Os06g0194000, Os07g0583700, Os03g0854500, and Os06g0105800, which are tagged by numbers 1, 2, 3, and 4 in Figure 6. The major regulatory paths are summarized in Table 1. We discuss regulation of the modules in detail in “Potential regulatory paths to the five functional modules” in the Discussion.

Multi-Omics Data Analysis of Differential Network Modules

The TF network analysis suggested that the five modules in the dehydration TF network were differentially expressed between WT and *erf71* in response to dehydration stress. To further investigate the regulation of the five modules in the dehydration TF network, we generated and analyzed the DNA methylome and miRNA sequencing data measured at 0, 1, and 6 HAT. DNA methylation sequencing data was used to investigate whether expression of genes in the five modules in the dehydration TF network was affected by DNA methylation, i.e., TF-DNA interaction, especially in the promoter regions. In addition, we also used miRNA sequencing data to investigate whether these genes were affected by miRNAs.

DNA Methylation Analysis to Investigate TF-DNA Interaction

Whole-genome bisulfite sequencing was performed at 0, 1, and 6 HAT in WT and *erf71* rice plants. Analysis of DNA methylation profiles showed that the average methylation levels in *erf71*



were slightly lower than those in WT at the whole-genome level. Meanwhile, differences in methylation level between time points were not seen (**Figure 7A**), which suggested that DNA methylation did not change over the time course of dehydration stress. We then investigated differences in the methylome between WT and *erf71* and found 1,607 DMRs with $p < e-12$ under Poison' distribution (see "Identifying differentially methylated regions" in Materials and methods). The DMRs covered 0.15% of the genome, and the majority of DMRs (1,436 DMRs) were hypomethylated in *erf71*. This result is consistent with studies observing that transgenic plants show stable hypomethylation on overall DNA compared with non-transgenic plants (Stroud et al., 2013; Stelpflug et al., 2014).

Searching for overlaps between promoters of genes and DMRs, we found that promoters of 494 genes (1.3% of total genes) overlapped with DMRs. Among them, 62 genes belonged to the five modules (2, 15, 5, 9, and 21 genes to Modules 1 to 5, respectively). From the Fisher's exact test, the hypothesis that there was a greater number of DMR-overlapping genes in the five modules vs. DMR-overlapping genes not in the five modules was not significant (i.e., $p > 0.05$), as shown in **Table 2**. This suggested that gene expression differences of the modules were not likely due to differences in DNA methylation.

miRNA Analysis to Show Potential miRNA Interference on Modules

In addition to TFs, another major mechanism of gene control is by miRNAs. To investigate whether miRNAs affected genes in the *OsERF71* cascade tree, small RNA sequencing was performed

at 0, 1, and 6 HAT in WT and *erf71* rice plants. At 1 and 6 HAT, differentially expressed miRNAs (DEmiRNAs) were identified (fold change >2 or fold change $<1/2$) (**Figure 7B**). There was one DEmiRNA during the 0-to-1 HAT period in both plants, and 20 and 15 DEmiRNAs during the 0-to-6 HAT period in WT and *erf71*, respectively. Among DEmiRNAs, we found that *osa-miR166j* was down-regulated during the 0-to-6 HAT period in both WT and *erf71*. Down-regulation of the MIR166 family in response to drought in rice was reported in previous studies (Zhou et al., 2010; Barrera-Figueroa et al., 2012; Cheah et al., 2015). As a next step, we performed a target gene analysis for the 32 unique DEmiRNAs and found 742 genes as potential targeted genes. Among them, 222 genes belonged to the five modules (4, 66, 10, 55, and 87 genes to Modules 1 to 5, respectively). The result of Fisher's exact test showed that DEmiRNA target genes were significantly represented in the modules (five modules and Modules 2, 4, and 5) compared with target genes not in the modules, as shown in **Table 3**. This suggested that gene expression differences in the modules were possibly due to regulation by miRNAs as well as TFs.

DISCUSSION

TF Network Construction

The TF network is the major computational resource for investigating biological mechanisms under dehydration conditions. Thus, we investigated all issues related to the TF network construction thoroughly.

TABLE 1 | The major regulatory paths to the five modules that were observed in the *OsERF71* cascade tree. Nine major regulatory paths that covered more than 10% of genes in the modules were observed in the *OsERF71* cascade tree.

No	Gene number	Gene ID	Gene descriptions	Coverage (% of Module)	The most significant GO term	P-value
1	1	Os06g0194000	<i>OsERF71</i>	66 % of Module 1	Translation	2.1e-55
	12	Os12g0597700	Similar to WRKY2			
2	1	Os06g0194000	<i>OsERF71</i>	21 % of Module 1	Translation	1.4e-10
	2	Os07g0583700	WRKY transcription factor 78			
	3	Os08g0157900	Similar to NAM protein			
3	1	Os06g0194000	<i>OsERF71</i>	56 % of Module 5	Photosynthesis	2.6e-21
	2	Os07g0583700	WRKY transcription factor 78			
	5	Os03g0854500	Similar to Heat shock transcription factor 31			
4	6	Os06g0105800	Homeodomain-like containing protein.	23 % of Module 537 % of Module 3	Citrate transport	3.9e-04
	1	Os06g0194000	<i>OsERF71</i>			
	2	Os07g0583700	WRKY transcription factor 78			
	5	Os03g0854500	Similar to Heat shock transcription factor 31			
5	8	Os03g0680800	Similar to cDNA clone: J013124I05	46 % of Module 2	Translation	7.9e-10
	1	Os06g0194000	<i>OsERF71</i>			
	16	Os06g0152200	Zinc-finger protein R2931			
6	17	Os11g0544700	Repressor protein.	35 % of Module 2	Translation	2.4e-100
	1	Os06g0194000	<i>OsERF71</i>			
	12	Os12g0597700	Similar to WRKY2			
	13	Os03g0318600	Similar to Transcription factor HBP-1b			
7	15	Os10g0561400	Similar to Transcription factor MYBS3.	51 % of Module 327 % of Module 4	Response to oxidative stress	1.9e-12
	1	Os06g0194000	<i>OsERF71</i>			
	2	Os07g0583700t	WRKY transcription factor 78			
	9	Os03g0795900	Similar to Heat shock transcription factor 31			
8	10	Os06g0140400	Similar to HAHB-6	19 % of Module 4	Lipid metabolic process	2.3e-05
	1	Os06g0194000	<i>OsERF71</i>			
	2	Os07g0583700	WRKY transcription factor 78			
	9	Os03g0795900	Similar to Heat shock transcription factor 31			
9	11	Os01g0211800	Similar to VirE2-interacting protein VIP1.	14 % of Module 4	DNA replication	2.3e-06
	1	Os06g0194000	<i>OsERF71</i>			
	12	Os12g0597700	Similar to WRKY2			
	13	Os03g0318600	Similar to Transcription factor HBP-1b			
	14	Os04g0676700	Similar to H0101F08.8 protein			

Gene numbers in the second column indicate the numbers given in **Figure 6**. The gene IDs of TFs in the paths are shown in the third column, where bold text indicates that the genes belong to Module 1. The most enriched GO term and its p-value for the downstream genes in each regulatory path are described in right-most column.

- Choice of the correlation-based TF network construction method.
- Effect of amount of omics data for network construction.
- Effect of cutoff values for a minimum correlation between genes.
- Comparison with other network construction and analysis methods.

Choice of the Correlation-Based TF Network Construction Method

Gene regulation network construction methods centered around TFs have been extensively studied over the years, and

network-based analysis of omics data has been successful in reconstructing gene regulatory paths under specific conditions. For example, Califano and colleagues (Basso et al., 2005) demonstrated a reverse-engineered construction of regulatory networks in human B-cells using gene expression data. A recent comprehensive study (Marbach et al., 2012) of biological network construction methods classified network construction methods into several groups based on the techniques used: regression (Haury et al., 2012), mutual information (Faith et al., 2007), correlation (Butte and Kohane, 2000), Bayesian networks (Statnikov and Aliferis, 2010), other approaches (Huynh-Thu et al., 2010), and meta predictors (Greenfield et al., 2010). According to the study (Marbach et al., 2012), correlation-based

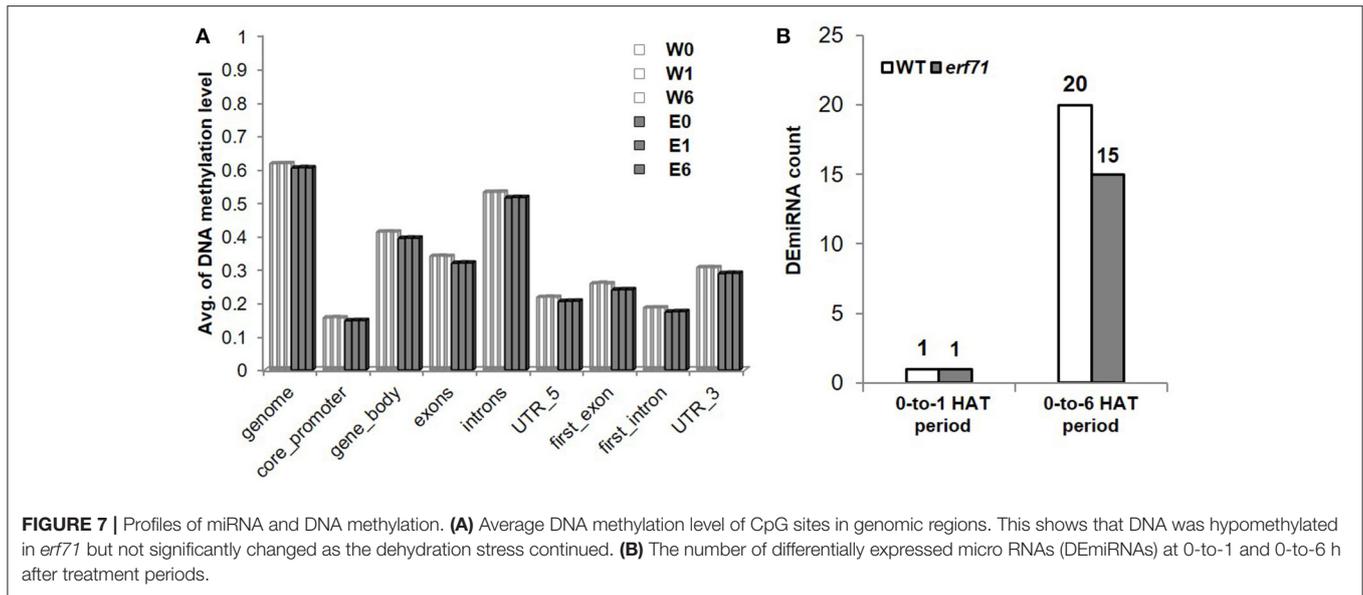


TABLE 2 | Results of the association test between differential DNA methylation and the five gene modules.

Gene module	DMR genes in module	Genes in module	DMR genes not in module	Genes not in module	<i>p</i> -value
Module 1	2	713	492	37,156	1.00
Module 2	15	1,363	479	36,506	0.78
Module 3	5	537	489	37,332	0.83
Module 4	19	1,586	475	36,283	0.68
Module 5	21	1,605	473	36,264	0.53

After identifying differentially methylated regions (DMRs), genes whose promoters overlapped with DMRs were investigated. Fisher's exact test was then used to evaluate the hypothesis that DMR-overlapping genes were observed more frequently in the modules compared with not in the modules. The results showed no significant association between differential DNA methylation and the five gene modules.

network construction is effective for investigating feed-forward networks. Our study on the role of *OsERF71* in the acquisition of the drought-resistant phenotype can be seen as investigating feed-forward networks originating from *OsERF71*. Therefore, we used a method based on Pearson's correlation for network construction.

Effect of Amount of Omics Data for Network Construction

In addition to the methods used for network construction, the amount of omics data used for network construction has a significant impact on the network topology. Hence, it was important to investigate whether the omics data we used for this study was sufficient to produce a network topology invariant of the size of data set used. To investigate this, we performed comprehensive network construction experiments with varying numbers of microarray data sets, from four to 1,893, doubling the microarray data size at each round. To reduce sampling bias,

TABLE 3 | Results of the association test between target genes of differentially expressed miRNAs (DEmiRNAs) and the five gene modules.

Gene module	DEmiRNA genes in module	Genes in module	DEmiRNA genes not in module	Genes not in module	<i>p</i> -value
Module 1	4	713	738	37,156	1.00
Module 2	66	1,363	676	36,506	2.0e-11
Module 3	10	537	732	37,332	0.61
Module 4	55	1,586	687	36,283	3.7e-05
Module 5	87	1,605	655	36,264	1.2e-17

After determining target genes of the DEmiRNAs, Fisher's exact test was used to evaluate the hypothesis that DEmiRNA target genes were observed more frequently in the modules compared with not in the modules. The results showed that the associations between DEmiRNA target genes and Modules 2, 3, and 5 were significant (Bold value indicates $p < 0.05$).

we used 10 randomly sampled microarray data sets for each network experiment with n samples from $n = 4$ to $n = 1,893$; i.e., the network was constructed 10 times given a sample size n . We then found that as the sample size increased, the Pearson's correlation coefficients converged into a normal-distribution-like shape (Supplemental Figure 3). As the sample size increased, the topology differences among the 10 networks with different samples disappeared, converging into a single network topology. The network topology was the same in more than 90% of network construction experiments from sample size 800 to 1,893 (Supplemental Figure 4).

Effect of Cutoff Values for a Minimum Correlation between Genes

The network topology is determined by a cutoff value for a minimum correlation value between two genes. To choose a reliable cutoff value, we measured the functionality scores as

defined in Materials and methods using varying cutoff values and then found that the cutoff value of 0.67 maximized the functionality score (Supplemental Figure 5). To investigate the effect of cutoff value, we generated networks using varying cutoff values. We observed that the five modules were consistently produced at PCC cutoff values ranging from 0.5 to 0.85 (Supplemental Figure 6). This result shows that the TF network module analysis generated robust results for varying cutoff values.

Comparison with Other Network Construction and Analysis Methods

Our choice of the correlation-based network construction method was based on a recent study by Marbach et al. (2012) that reported that correlation-based network construction is the most effective method for recovering feed-forward loops, which can be seen as a regulation mechanism of the feed-forward propagation from *OsERF71*, which was the subject of this study. However, we also examined two major network construction and analysis methods as described below.

We constructed another template TF network using a state-of-the-art gene-expression-based TF network construction method, Narromi (Zhang et al., 2013). Using 1,893 microarray data sets, Narromi produced 6,213,143 TF-target gene edges ($p < 0.05$). Edges of strong association ($|PCC| > 0.67$) in eight mRNA-seq data sets were then selected. The Narromi-based dehydration TF network consisted of 23,236 genes and 1,286,285 edges. Of these, 7,314 genes, and 28,616 edges were common with our correlation-based network. In the network, seven gene cluster structures were observed (Supplemental Figure 7A). Among the clusters, we found that five gene clusters were differentially expressed (Supplemental Table 2). In summary, the discovery of five functional modules was confirmed by the network constructed using Narromi.

Another TF network analysis was performed using RiceNet (Lee et al., 2011), a well-curated gene-gene interaction network. RiceNet consists of 14,949 TF-target gene edges. Since our goal was to characterize the role of a knockout TF, we excluded edges between nonTFs. Edges with strong association ($|PCC| > 0.67$) in eight mRNA-seq data sets were then selected. The RiceNet-induced dehydration TF network consisted of 2,345 genes and 3,667 edges. Of these, 1,234 genes and 456 edges were common with our correlation-based network. In the network, 137 gene cluster structures were observed (Supplemental Figure 7B). Among the clusters, we found that one gene cluster was differentially expressed (Supplemental Table 2). Only one network module was detected in the RiceNet network. We conjecture that this is because we excluded edges between two nonTF genes, which could affect the formation of gene clusters. In addition, RiceNet is a general template network that is not designed for condition-specific gene expression data.

Characteristic Physiological Mechanisms in Modules

Many genes in Module 5 were related to photosynthesis and energy-generating mechanisms: functions of genes

were related to light (e.g., Os01g0764500, similar to *uvrB/uvrC* motif-containing protein) and photosynthesis (e.g., Os01g0773700, similar to photosystem II reaction center W protein; Os03g0267300, similar to fructose-1,6-bisphosphatase; Os12g0291100, Rubisco small subunit), energy transfer (e.g., Os03g0278900, ATPase; Os11g0661300, similar to ADP/ATP translocase-like protein), and subsequent anabolic pathways, such as those for carbohydrate (e.g., Os01g0686200, UDP-glucuronosyl/UDP-glucosyltransferase family protein), carbohydrate movement and translocation (e.g., Os03g0363500, similar to vacuolar monosaccharide symporter 1), amino acid synthesis (nitrogen fixation) (e.g., Os06g0694500, similar to nitrogen fixation like protein; Os07g0658400, similar to ferredoxin-dependent glutamate synthase), and lipid synthesis (e.g., Os02g0589000, lecithin:cholesterol acyltransferase family protein). In summary, based on functions of the genes in Module 5, biological mechanisms of Module 5 were related to energy generation, storage and transfer. Recall that genes in Module 5 were more down-regulated in *erf71* than in WT at 1 HAT. This suggests a fast and flexible response in the early stages of drought stress that may be directly and/or indirectly related to the successive physiological mechanisms observed in Modules 2, 3, and 4.

Discontinuation of water supply to the plant is known to result in loss of tension and decrease in water potential in the cells of leaves, which directly affects most metabolic pathways, particularly photosynthesis as shown above. The widely known water split process that activates an electron and releases it from the water molecules in the reaction center of PSII is affected by water shortage in photosynthetic cells. Most of the genes that code for proteins of PSII were down-regulated in both WT and *erf71*, which is consistent with our current knowledge.

ROS, including hydrogen peroxide (H_2O_2), superoxide anion (O_2^-), hydroxyl radical ($\cdot OH$) and singlet oxygen (1O_2), are produced in chloroplasts, mitochondria, peroxisomes, cell membranes and cell wall spaces (Moradi and Ismail, 2007). Many enzymes and physical processes are involved in ROS production and scavenging. A large portion of ROS is produced by NADPH oxidase (NOX) during photosynthesis and photorespiration, and representative ROS scavenging enzymes are superoxide dismutase (SOD), catalase (CAT), peroxidase (Perox), and ascorbate peroxidase (APX) while non-enzymatic scavenging systems include various flavonoids, alkaloids, phenolic compounds, tocopherols, carotenoids, and metallothioneins (MTs) (Ueda et al., 2013; Han et al., 2014). It has traditionally been believed that under conditions of limited water supply the response of rice plants is to overcome physical and chemical damage and that the generation of ROS is the result of accumulation from cellular breakdown. However, recent studies have revealed that ROS may confer protection against water stress (Kar, 2011). ROS may be involved in homeostasis and various metabolic changes, morphological and anatomical changes, and metabolic adaptation by rice plants under water stress closely related to drought avoidance or postponement and alleviation of stress-induced cellular injuries.

Potential Regulatory Paths to the Five Functional Modules

TF network analysis showed that Module 1 was associated with drought-stress-related TFs. That is, 20% of genes in Module 1 consisted of TFs, most of which belonged to water-stress-related TF families. In addition, only Module 1 was up-regulated in both WT and *erf71* in response to dehydration stress whereas the other modules were down-regulated. Overexpressing *OsERF71* influenced Module 1 by reducing the up-regulation of expression of the member genes under dehydration stress. Analysis of an *OsERF71* cascade tree showed that TFs in Module 1 were present in every upstream of nine major regulatory paths as highlighted in bold type in **Table 1**. These results suggest that the down-regulation of overall gene expression is initiated upon up-regulation of the stress-induced TF module. We also suggest nine major regulatory paths determining how the downstream genes of the modules are regulated during dehydration stress treatment. However, since these were established by inference, they should be validated by further experiments. In addition, we found that miRNA was also involved in the control of gene expression of the five differentially expressed modules. For instance, we found that *osa-miR319a* was up-regulated under dehydration stress in *erf71*, which is consistent with the observations of a previous study (Zhou et al., 2010). The 8, 12, 2, and 5 predicted target genes of *osa-miR319a* were distributed in Modules 2 to 5, respectively. Regulation by both TF and miRNA as we suggest is statistically supported, but there needs to be experimental validation for each path and thus future study is required.

ACCESSION NUMBERS

Raw sequencing data are available via accession ID: GSE74465 in the Gene Expression Omnibus (GEO) database.

REFERENCES

- Ambavaram, M. M., Basu, S., Krishnan, A., Ramegowda, V., Batlang, U., Rahman, L., et al. (2014). Coordinated regulation of photosynthesis in rice increases yield and tolerance to environmental stress. *Nat. Commun.* 5:5302. doi: 10.1038/ncomms6302
- Barabasi, A. L., Gulbahce, N., and Loscalzo, J. (2011). Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* 12, 56–68. doi: 10.1038/nrg2918
- Barabasi, A. L., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* 5, 101–113. doi: 10.1038/nrg1272
- Barrera-Figueroa, B. E., Gao, L., Wu, Z., Zhou, X., Zhu, J., Jin, H., et al. (2012). High throughput sequencing reveals novel and abiotic stress-regulated microRNAs in the inflorescences of rice. *BMC Plant Biol.* 12:132. doi: 10.1186/1471-2229-12-132
- Basso, K., Margolin, A. A., Stolovitzky, G., Klein, U., Dalla-Favera, R., and Califano, A. (2005). Reverse engineering of regulatory networks in human B cells. *Nat. Genet.* 37, 382–390. doi: 10.1038/ng1532
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* 2008:P10008. doi: 10.1088/1742-5468/2008/10/P10008
- Boutillier, K., Offringa, R., Sharma, V. K., Kieft, H., Ouellet, T., Zhang, L., et al. (2002). Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* 14, 1737–1749. doi: 10.1105/tpc.001941

AUTHOR CONTRIBUTIONS

JK provided *OsERF71* transgenic rice seeds. SK and HK designed the project. HK and SS designed and conducted the drought experiments. HA and IJ processed the omics data and performed bioinformatics analysis. HA and JP performed network analysis, and SR performed miRNA analysis. WJ interpreted the network analysis results biologically. HA, IJ, SK, and WJ wrote the paper.

FUNDING

This work was conducted with the support of Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ01121102), Rural Development Administration, Republic of Korea, of Collaborative Genome Program for Fostering New Post-Genome industry through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2014M3C9A3063541), of BK21 Plus for Pioneers in Innovative Computing (Dept. of Computer Science and Engineering, SNU) funded by National Research Foundation of Korea (NRF) (21A20151113068), and of the Agenda program (No. PJ01142502), Rural Development Administration, Republic of Korea.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpls.2017.01044/full#supplementary-material>

- Butte, A. J., and Kohane, I. S. (2000). Mutual information relevance networks: functional genomic clustering using pairwise entropy measurements. *Pac. Symp. Biocomput.* 5, 418–429. doi: 10.1142/9789814447331_0040
- Cheah, B. H., Nadarajah, K., Divate, M. D., and Wickneswari, R. (2015). Identification of four functionally important microRNA families with contrasting differential expression profiles between drought-tolerant and susceptible rice leaf at vegetative stage. *BMC Genomics* 16:692. doi: 10.1186/s12864-015-1851-3
- Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., and Yu, D. (2012). The role of WRKY transcription factors in plant abiotic stresses. *Biochim. Biophys. Acta* 1819, 120–128. doi: 10.1016/j.bbagr.2011.09.002
- Chen, W., Provart, N. J., Glazebrook, J., Katagiri, F., Chang, H. S., Eulgem, T., et al. (2002). Expression profile matrix of Arabidopsis transcription factor genes suggests their putative functions in response to environmental stresses. *Plant Cell* 14, 559–574. doi: 10.1105/tpc.010410
- Chi, Y. Y., Gribbin, M. J., Johnson, J. L., and Muller, K. E. (2014). Power calculation for overall hypothesis testing with high-dimensional commensurate outcomes. *Stat. Med.* 33, 812–827. doi: 10.1002/sim.5986
- Chodavarapu, R. K., Feng, S., Ding, B., Simon, S. A., Lopez, D., Jia, Y., et al. (2012). Transcriptome and methylome interactions in rice hybrids. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12040–12045. doi: 10.1073/pnas.1209297109
- Chuck, G., Meeley, R. B., and Hake, S. (1998). The control of maize spikelet meristem fate by the APETALA2-like gene indeterminate spikelet1. *Genes Dev.* 12, 1145–1154. doi: 10.1101/gad.12.8.1145

- Dai, X., and Zhao, P. X. (2011). psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Res.* 39, W155–W159. doi: 10.1093/nar/gkr319
- Debnath, M., Pandey, M., and Bisen, P. S. (2011). An omics approach to understand the plant abiotic stress. *OMICS* 15, 739–762. doi: 10.1089/omi.2010.0146
- Degenkolbe, T., Do, P. T., Zuther, E., Reipsilber, D., Walther, D., Hinch, D. K., et al. (2009). Expression profiling of rice cultivars differing in their tolerance to long-term drought stress. *Plant Mol. Biol.* 69, 133–153. doi: 10.1007/s11103-008-9412-7
- Du, H., Wang, N., Cui, F., Li, X., Xiao, J., and Xiong, L. (2010). Characterization of the β -carotene hydroxylase gene DSM2 conferring drought and oxidative stress resistance by increasing xanthophylls and abscisic acid synthesis in rice. *Plant Physiol.* 154, 1304–1318. doi: 10.1104/pp.110.163741
- Dubouzet, J. G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E. G., Miura, S., et al. (2003). OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *Plant J.* 33, 751–763. doi: 10.1046/j.1365-3113X.2003.01661.x
- Elliott, R. C., Betzner, A. S., Huttner, E., Oakes, M. P., Tucker, W. Q., Gerentes, D., et al. (1996). AINTEGUMENTA, an APETALA2-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth. *Plant Cell* 8, 155–168. doi: 10.1105/tpc.8.2.155
- Faith, J. J., Hayete, B., Thaden, J. T., Mogno, I., Wierzbowski, J., Cottarel, G., et al. (2007). Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* 5:e8. doi: 10.1371/journal.pbio.0050008
- Fang, Y., You, J., Xie, K., Xie, W., and Xiong, L. (2008). Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice. *Mol. Genet. Genomics* 280, 547–563. doi: 10.1007/s00438-008-0386-6
- Ficklin, S. P., and Feltus, F. A. (2011). Gene coexpression network alignment and conservation of gene modules between two grass species: maize and rice. *Plant Physiol.* 156, 1244–1256. doi: 10.1104/pp.111.173047
- Fowlkes, E. B., and Mallows, C. L. (1983). A Method for comparing two hierarchical clusterings. *J. Am. Stat. Assoc.* 78, 553–569. doi: 10.1080/01621459.1983.10478008
- Ge, L. F., Chao, D. Y., Shi, M., Zhu, M. Z., Gao, J. P., and Lin, H. X. (2008). Overexpression of the trehalose-6-phosphate phosphatase gene OsTPP1 confers stress tolerance in rice and results in the activation of stress responsive genes. *Planta* 228, 191–201. doi: 10.1007/s00425-008-0729-x
- Gill, S. S., and Tuteja, N. (2010). Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiol. Biochem.* 48, 909–930. doi: 10.1016/j.plaphy.2010.08.016
- Gitter, A., Carmi, M., Barkai, N., and Bar-Joseph, Z. (2013). Linking the signaling cascades and dynamic regulatory networks controlling stress responses. *Genome Res.* 23, 365–376. doi: 10.1101/gr.138628.112
- Gorantla, M., Babu, P. R., Lachagari, V. B., Reddy, A. M., Wusirika, R., Bennetzen, J. L., et al. (2007). Identification of stress-responsive genes in an indica rice (*Oryza sativa* L.) using ESTs generated from drought-stressed seedlings. *J. Exp. Bot.* 58, 253–265. doi: 10.1093/jxb/erl213
- Greenfield, A., Madar, A., Ostrer, H., and Bonneau, R. (2010). DREAM4: combining genetic and dynamic information to identify biological networks and dynamical models. *PLoS ONE* 5:e13397. doi: 10.1371/journal.pone.0013397
- Guo, L., Wang, Z. Y., Lin, H., Cui, W. E., Chen, J., Liu, M., et al. (2006). Expression and functional analysis of the rice plasma-membrane intrinsic protein gene family. *Cell Res.* 16, 277–286. doi: 10.1038/sj.cr.7310035
- Haake, V., Cook, D., Riechmann, J. L., Pineda, O., Thomashow, M. F., and Zhang, J. Z. (2002). Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. *Plant Physiol.* 130, 639–648. doi: 10.1104/pp.006478
- Hackenbreg, M., Rodriguez-Ezpeleta, N., and Aransay, A. M. (2011). miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments. *Nucleic Acids Res.* 39, W132–W138. doi: 10.1093/nar/gkr247
- Hadiarto, T., and Tran, L. S. (2011). Progress studies of drought-responsive genes in rice. *Plant Cell Rep.* 30, 297–310. doi: 10.1007/s00299-010-0956-z
- Han, M., Kim, C. Y., Lee, J., Lee, S. K., and Jeon, J. S. (2014). OsWRKY42 represses OsMT1d and induces reactive oxygen species and leaf senescence in rice. *Mol. Cells* 37, 532–539. doi: 10.14348/molcells.2014.0128
- Haury, A. C., Mordelet, F., Vera-Licona, P., and Vert, J. P. (2012). TIGRESS: Trustful Inference of Gene REgulation using Stability Selection. *BMC Syst. Biol.* 6:145. doi: 10.1186/1752-0509-6-145
- Hsieh, T. F., Ibarra, C. A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R. L., et al. (2009). Genome-wide demethylation of Arabidopsis endosperm. *Science* 324, 1451–1454. doi: 10.1126/science.1172417
- Hsieh, T. H., Lee, J. T., Yang, P. T., Chiu, L. H., Charng, Y. Y., Wang, Y. C., et al. (2002). Heterology expression of the Arabidopsis C-repeat/dehydration response element binding factor 1 gene confers elevated tolerance to chilling and oxidative stresses in transgenic tomato. *Plant Physiol.* 129, 1086–1094. doi: 10.1104/pp.003442
- Hu, H., Dai, M., Yao, J., Xiao, B., Li, X., Zhang, Q., et al. (2006). Overexpressing a NAM, ATAF, and CUC (NAC) transcription factor enhances drought resistance and salt tolerance in rice. *Proc. Natl. Acad. Sci. U.S.A.* 103, 12987–12992. doi: 10.1073/pnas.0604882103
- Huang, X. Y., Chao, D. Y., Gao, J. P., Zhu, M. Z., Shi, M., and Lin, H. X. (2009). A previously unknown zinc finger protein, DST, regulates drought and salt tolerance in rice via stomatal aperture control. *Genes Dev.* 23, 1805–1817. doi: 10.1101/gad.1812409
- Huynh-Thu, V. A., Irrthum, A., Wehenkel, L., and Geurts, P. (2010). Inferring regulatory networks from expression data using tree-based methods. *PLoS ONE* 5:e12776. doi: 10.1371/journal.pone.0012776
- Hwang, S. G., Kim, D. S., Hwang, J. E., Han, A. R., and Jang, C. S. (2014). Identification of rice genes associated with cosmic-ray response via co-expression gene network analysis. *Gene* 541, 82–91. doi: 10.1016/j.gene.2014.02.060
- IPCC (2014). “Climate Change 2014: Impacts, Adaptation, and Vulnerability,” in Part B: Regional Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge; New York, NY: Cambridge University Press.
- Islam, M. A., Du, H., Ning, J., Ye, H., and Xiong, L. (2009). Characterization of Glossy1-homologous genes in rice involved in leaf wax accumulation and drought resistance. *Plant Mol. Biol.* 70, 443–456. doi: 10.1007/s11103-009-9483-0
- Iuchi, S., Kobayashi, M., Taji, T., Naramoto, M., Seki, M., Kato, T., et al. (2001). Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in Arabidopsis. *Plant J.* 27, 325–333. doi: 10.1046/j.1365-3113x.2001.01096.x
- Jeong, J. S., Kim, Y. S., Baek, K. H., Jung, H., Ha, S. H., Do Choi, Y., et al. (2010). Root-specific expression of OsNAC10 improves drought tolerance and grain yield in rice under field drought conditions. *Plant Physiol.* 153, 185–197. doi: 10.1104/pp.110.154773
- Jin, J., Zhang, H., Kong, L., Gao, G., and Luo, J. (2014). PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.* 42, D1182–D1187. doi: 10.1093/nar/gkt1016
- Kar, R. K. (2011). Plant responses to water stress: role of reactive oxygen species. *Plant Signal. Behav.* 6, 1741–1745. doi: 10.4161/psb.6.11.17729
- Kawahara, Y., De La Bastide, M., Hamilton, J. P., Kanamori, H., Mccombie, W. R., Ouyang, S., et al. (2013). Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. *Rice (N. Y.)* 6:4. doi: 10.1186/1939-8433-6-4
- Kozomara, A., and Griffiths-Jones, S. (2014). miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res.* 42, D68–D73. doi: 10.1093/nar/gkt1181
- Lee, D. K., Jung, H., Jang, G., Jeong, J. S., Kim, Y. S., Ha, S. H., et al. (2016). Overexpression of the OsERF71 transcription factor alters rice root structure and drought resistance. *Plant Physiol.* 172, 575–588. doi: 10.1104/pp.16.00379
- Lee, I., Seo, Y. S., Coltrane, D., Hwang, S., Oh, T., Marcotte, E. M., et al. (2011). Genetic dissection of the biotic stress response using a genome-scale gene network for rice. *Proc. Natl. Acad. Sci. U.S.A.* 108, 18548–18553. doi: 10.1073/pnas.1110384108
- Lian, H. L., Yu, X., Ye, Q., Ding, X., Kitagawa, Y., Kwak, S. S., et al. (2004). The role of aquaporin RWC3 in drought avoidance in rice. *Plant Cell Physiol.* 45, 481–489. doi: 10.1093/pcp/pch058
- Lim, S. D., Yim, W. C., Moon, J. C., Kim, D. S., Lee, B. M., and Jang, C. S. (2010). A gene family encoding RING finger proteins in rice: their expansion, expression diversity, and co-expression genes. *Plant Mol. Biol.* 72, 369–380. doi: 10.1007/s11103-009-9576-9

- Lindemose, S., O'shea, C., Jensen, M. K., and Skriver, K. (2013). Structure, function and networks of transcription factors involved in abiotic stress responses. *Int. J. Mol. Sci.* 14, 5842–5878. doi: 10.3390/ijms14035842
- Lu, Z., Huang, X., Ouyang, Y., and Yao, J. (2013). Genome-wide identification, phylogenetic and co-expression analysis of OsSET gene family in rice. *PLoS ONE* 8:e65426. doi: 10.1371/journal.pone.0065426
- Ma, C., Xin, M., Feldmann, K. A., and Wang, X. (2014). Machine learning-based differential network analysis: a study of stress-responsive transcriptomes in Arabidopsis. *Plant Cell* 26, 520–537. doi: 10.1105/tpc.113.121913
- Mao, L., Van Hemert, J. L., Dash, S., and Dickerson, J. A. (2009). Arabidopsis gene co-expression network and its functional modules. *BMC Bioinformatics* 10:346. doi: 10.1186/1471-2105-10-346
- Marbach, D., Costello, J. C., Kuffner, R., Vega, N. M., Prill, R. J., Camacho, D. M., et al. (2012). Wisdom of crowds for robust gene network inference. *Nat. Methods* 9, 796–804. doi: 10.1038/nmeth.2016
- Miller, J. A., Cai, C., Langfelder, P., Geschwind, D. H., Kurian, S. M., Salomon, D. R., et al. (2011). Strategies for aggregating gene expression data: the collapseRows R function. *BMC Bioinformatics* 12:322. doi: 10.1186/1471-2105-12-322
- Miyamoto, Y., Koh, Y. H., Park, Y. S., Fujiwara, N., Sakiyama, H., Misonou, Y., et al. (2003). Oxidative stress caused by inactivation of glutathione peroxidase and adaptive responses. *Biol. Chem.* 384, 567–574. doi: 10.1515/BC.2003.064
- Mizoi, J., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2012). AP2/ERF family transcription factors in plant abiotic stress responses. *Biochim. Biophys. Acta* 1819, 86–96. doi: 10.1016/j.bbagr.2011.08.004
- Moose, S. P., and Sisco, P. H. (1996). Glossy15, an APETALA2-like gene from maize that regulates leaf epidermal cell identity. *Genes Dev.* 10, 3018–3027. doi: 10.1101/gad.10.23.3018
- Moradi, F., and Ismail, A. M. (2007). Responses of photosynthesis, chlorophyll fluorescence and ROS-scavenging systems to salt stress during seedling and reproductive stages in rice. *Ann. Bot.* 99, 1161–1173. doi: 10.1093/aob/mcm052
- Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiol.* 140, 411–432. doi: 10.1104/pp.105.073783
- Nakashima, K., Takasaki, H., Mizoi, J., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2012). NAC transcription factors in plant abiotic stress responses. *Biochim. Biophys. Acta* 1819, 97–103. doi: 10.1016/j.bbagr.2011.10.005
- Nakashima, K., Yamaguchi-Shinozaki, K., and Shinozaki, K. (2014). The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. *Front. Plant Sci.* 5:170. doi: 10.3389/fpls.2014.00170
- Oh, S. J., Kim, Y. S., Kwon, C. W., Park, H. K., Jeong, J. S., and Kim, J. K. (2009). Overexpression of the transcription factor AP37 in rice improves grain yield under drought conditions. *Plant Physiol.* 150, 1368–1379. doi: 10.1104/pp.109.137554
- Oh, S. J., Song, S. I., Kim, Y. S., Jang, H. J., Kim, S. Y., Kim, M., et al. (2005). Arabidopsis CBF3/DREB1A and ABF3 in transgenic rice increased tolerance to abiotic stress without stunting growth. *Plant Physiol.* 138, 341–351. doi: 10.1104/pp.104.059147
- Pellegrineschi, A., Reynolds, M., Pacheco, M., Brito, R. M., Almeraya, R., Yamaguchi-Shinozaki, K., et al. (2004). Stress-induced expression in wheat of the Arabidopsis thaliana DREB1A gene delays water stress symptoms under greenhouse conditions. *Genome* 47, 493–500. doi: 10.1139/g03-140
- Qin, X., and Zeevaert, J. A. (2002). Overexpression of a 9-cis-epoxycarotenoid dioxygenase gene in *Nicotiana glauca* increases abscisic acid and phaseic acid levels and enhances drought tolerance. *Plant Physiol.* 128, 544–551. doi: 10.1104/pp.010663
- Qiu, Y., and Yu, D. (2009). Over-expression of the stress-induced OsWRKY45 enhances disease resistance and drought tolerance in Arabidopsis. *Environ. Exp. Bot.* 65, 35–47. doi: 10.1016/j.envexpbot.2008.07.002
- Rabbani, M. A., Maruyama, K., Abe, H., Khan, M. A., Katsura, K., Ito, Y., et al. (2003). Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analyses. *Plant Physiol.* 133, 1755–1767. doi: 10.1104/pp.103.025742
- Rabello, A. R., Guimaraes, C. M., Rangel, P. H., Da Silva, F. R., Seixas, D., De Souza, E., et al. (2008). Identification of drought-responsive genes in roots of upland rice (*Oryza sativa* L.). *BMC Genomics* 9:485. doi: 10.1186/1471-2164-9-485
- Ramachandra Reddy, A., Chaitanya, K. V., and Vivekanandan, M. (2004). Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. *J. Plant Physiol.* 161, 1189–1202. doi: 10.1016/j.jplph.2004.01.013
- Rodriguez, M., Canales, E., Borroto, C. J., Carmona, E., Lopez, J., Pujol, M., et al. (2006). Identification of genes induced upon water-deficit stress in a drought-tolerant rice cultivar. *J. Plant Physiol.* 163, 577–584. doi: 10.1016/j.jplph.2005.07.005
- Saito, R., Smoot, M. E., Ono, K., Ruschinski, J., Wang, P. L., Lotia, S., et al. (2012). A travel guide to Cytoscape plugins. *Nat. Methods* 9, 1069–1076. doi: 10.1038/nmeth.2212
- Sakai, H., Lee, S. S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., et al. (2013). Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol.* 54:e6. doi: 10.1093/pcp/pcs183
- Seo, J. S., Joo, J., Kim, M. J., Kim, Y. K., Nahm, B. H., Song, S. I., et al. (2011). OsBHLH148, a basic helix-loop-helix protein, interacts with OsJAZ proteins in a jasmonate signaling pathway leading to drought tolerance in rice. *Plant J.* 65, 907–921. doi: 10.1111/j.1365-3113.2010.04477.x
- Shinozaki, K., and Yamaguchi-Shinozaki, K. (2007). Gene networks involved in drought stress response and tolerance. *J. Exp. Bot.* 58, 221–227. doi: 10.1093/jxb/erl164
- Shinozaki, K., Yamaguchi-Shinozaki, K., and Seki, M. (2003). Regulatory network of gene expression in the drought and cold stress responses. *Curr. Opin. Plant Biol.* 6, 410–417. doi: 10.1016/S1369-5266(03)00092-X
- Shojaie, A., and Michailidis, G. (2009). Analysis of gene sets based on the underlying regulatory network. *J. Comput. Biol.* 16, 407–426. doi: 10.1089/cmb.2008.0081
- Singh, D., and Laxmi, A. (2015). Transcriptional regulation of drought response: a tortuous network of transcriptional factors. *Front. Plant Sci.* 6:895. doi: 10.3389/fpls.2015.00895
- Song, S. Y., Chen, Y., Chen, J., Dai, X. Y., and Zhang, W. H. (2011). Physiological mechanisms underlying OsNAC5-dependent tolerance of rice plants to abiotic stress. *Planta* 234, 331–345. doi: 10.1007/s00425-011-1403-2
- Statnikov, A., and Aliferis, C. F. (2010). Analysis and computational dissection of molecular signature multiplicity. *PLoS Comput. Biol.* 6:e1000790. doi: 10.1371/journal.pcbi.1000790
- Stelpflug, S. C., Eichten, S. R., Hermanson, P. J., Springer, N. M., and Kaeppler, S. M. (2014). Consistent and heritable alterations of DNA methylation are induced by tissue culture in maize. *Genetics* 198, 209–218. doi: 10.1534/genetics.114.165480
- Stroud, H., Ding, B., Simon, S. A., Feng, S., Bellizzi, M., Pellegrini, M., et al. (2013). Plants regenerated from tissue culture contain stable epigenome changes in rice. *Elife* 2:e00354. doi: 10.7554/eLife.00354
- The Mutation and Pathway Analysis Working Group of the International Cancer Genome (2015). Pathway and network analysis of cancer genomes. *Nat. Meth.* 12, 615–621. doi: 10.1038/nmeth.3440
- Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111. doi: 10.1093/bioinformatics/btp120
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621
- Ueda, Y., Uehara, N., Sasaki, H., Kobayashi, K., and Yamakawa, T. (2013). Impacts of acute ozone stress on superoxide dismutase (SOD) expression and reactive oxygen species (ROS) formation in rice leaves. *Plant Physiol. Biochem.* 70, 396–402. doi: 10.1016/j.plaphy.2013.06.009
- Vanhohle, B., Grunewald, W., Bateman, A., Kohchi, T., and Gheysen, G. (2007). The tify family previously known as ZIM. *Trends Plant Sci.* 12, 239–244. doi: 10.1016/j.tplants.2007.04.004
- Venuprasad, R., Lafitte, H. R., and Atlin, G. N. (2007). Response to direct selection for grain yield under drought stress in rice this research work was funded by BMZ, Germany, and the Rockefeller Foundation, USA. *Crop Sci.* 47, 285–293. doi: 10.2135/cropsci2006.03.0181
- Wang, Q., Guan, Y., Wu, Y., Chen, H., Chen, F., and Chu, C. (2008). Overexpression of a rice OsDREB1F gene increases salt, drought, and low temperature tolerance in both Arabidopsis and rice. *Plant Mol. Biol.* 67, 589–602. doi: 10.1007/s11103-008-9340-6

- Xi, Y., and Li, W. (2009). BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* 10:232. doi: 10.1186/1471-2105-10-232
- Xu, D. Q., Huang, J., Guo, S. Q., Yang, X., Bao, Y. M., Tang, H. J., et al. (2008). Overexpression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (*Oryza sativa* L.). *FEBS Lett.* 582, 1037–1043. doi: 10.1016/j.febslet.2008.02.052
- Yang, Y., Zhong, J., Ouyang, Y. D., and Yao, J. (2013). The integrative expression and co-expression analysis of the AGO gene family in rice. *Gene* 528, 221–235. doi: 10.1016/j.gene.2013.07.002
- Yi, N., Kim, Y. S., Jeong, M. H., Oh, S. J., Jeong, J. S., Park, S. H., et al. (2010). Functional analysis of six drought-inducible promoters in transgenic rice plants throughout all stages of plant growth. *Planta* 232, 743–754. doi: 10.1007/s00425-010-1212-z
- Yoshida, S., Forno, D., and Cock, J. (1976). *Laboratory Manual for Physiological Studies of Rice*. Los Baños: International Rice Research Institute (IRRI).
- Zhang, H., Ni, L., Liu, Y., Wang, Y., Zhang, A., Tan, M., et al. (2012). The C2H2-type zinc finger protein ZFP182 is involved in abscisic acid-induced antioxidant defense in rice. *J. Integr. Plant Biol.* 54, 500–510. doi: 10.1111/j.1744-7909.2012.01135.x
- Zhang, L., Yu, S., Zuo, K., Luo, L., and Tang, K. (2012). Identification of gene modules associated with drought response in rice by network-based analysis. *PLoS ONE* 7:e33748. doi: 10.1371/journal.pone.0033748
- Zhang, X., Liu, K., Liu, Z. P., Duval, B., Richer, J. M., Zhao, X. M., et al. (2013). NARROMI: a noise and redundancy reduction technique improves accuracy of gene regulatory network inference. *Bioinformatics* 29, 106–113. doi: 10.1093/bioinformatics/bts619
- Zheng, X., Chen, B., Lu, G., and Han, B. (2009). Overexpression of a NAC transcription factor enhances rice drought and salt tolerance. *Biochem. Biophys. Res. Commun.* 379, 985–989. doi: 10.1016/j.bbrc.2008.12.163
- Zhou, J., Wang, X., Jiao, Y., Qin, Y., Liu, X., He, K., et al. (2007). Global genome expression analysis of rice in response to drought and high-salinity stresses in shoot, flag leaf, and panicle. *Plant Mol. Biol.* 63, 591–608. doi: 10.1007/s11103-006-9111-1
- Zhou, L., Liu, Y., Liu, Z., Kong, D., Duan, M., and Luo, L. (2010). Genome-wide identification and analysis of drought-responsive microRNAs in *Oryza sativa*. *J. Exp. Bot.* 61, 4157–4168. doi: 10.1093/jxb/erq237

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Ahn, Jung, Shin, Park, Rhee, Kim, Jung, Kwon and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.