



Genome-Wide Identification of Drought Response Genes in Soybean Seedlings and Development of Biomarkers for Early Diagnoses

Lei Hua¹ · Ghana Shyam Challa¹ · Senthil Subramanian^{1,2} · Xingyou Gu^{1,2} · Wanlong Li^{1,2} 

Published online: 11 May 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

Drought is a yield-limiting factor for soybean (*Glycine max* L. Merrill) production in North America. Understanding molecular mechanisms underlying early responses of the soybean plant to drought may help develop new techniques to manage the abiotic stress. The objectives of this research were to identify expressed genes responsive to drought stress at the seedling stage and to develop biomarkers for early diagnostic of genotypic difference in the stress tolerance. Using a GeneChip Soybean Genome Array and an improved algorithm, we identified 697 differentially expressed genes (DEGs), with 420 upregulated and 277 downregulated by a 6-h dehydration treatment. A majority of these DEGs encode transcription factors, protein kinases, hormone biosynthetic/signaling enzymes, or other regulatory proteins. MAPMAN and KEGG enrichment showed that the DEGs were mainly involved in the metabolic and hormone signaling pathways and identified the GmHDZ72/PYL/PP2C module as negative feedback of abscisic acid signaling pathway induced by dehydration stress. Ten DEGs were selected from various pathways and validated in a sample of 20 soybean cultivars varying in the level of drought tolerance. Five of the 10 validated DEGs, *Glyma03g30040*, *Glyma11g05960*, *Glyma11g11430*, *Glyma12g22880*, and *Glyma16g02390*, showed expression profiles strongly correlated with the plant height reduction after a 14-d drought treatment. These genes are the best candidate biomarkers to monitor plant early responses to drought stress before a symptom appears and to screen for drought-tolerant genotypes. This research provided a new set of transcriptomic data to develop gene regulatory networks underlying sensing drought signal and possible acclimatization in the early stage and enriched the genomics toolbox with a set of biomarkers for early diagnosis of drought damage and molecular breeding of drought tolerance in soybean.

Keywords Drought · Microarray · qRT-PCR · Biomarkers · Soybean

Introduction

As one of the most common environmental constraints affecting plant growth and yield, drought has long been a prevailing threat to crop production. Recent climate variability has increased the extent and frequency of drought damage. Development of drought-resistant crops

is a practical solution to this abiotic stress. As one of the most complex traits, variations of drought tolerance are controlled by many quantitative trait loci (QTL) with low heredity and further complicated by the interactions between genotype and environments. The genetic and molecular complexity significantly impedes breeding drought tolerance by marker-assisted selection (Chai et al. 2017; Collins et al. 2008; Shivhare and Lata 2017; Valliyodan et al. 2016). The recent development of more radical approaches to improve drought tolerance would be the application of contemporary genome-editing technologies such as CRISPR (Cong et al. 2013), which has been successful in the improvement of many other agronomic traits (reviewed by Weeks et al. 2016). One of the prerequisites of this approach is the identification of the major players and genetic pathways underlying the plant response to drought stress. In another aspect, information on early plant

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11105-018-1085-z>) contains supplementary material, which is available to authorized users.

✉ Wanlong Li
Wanlong.li@sdstate.edu

¹ Department of Biology and Microbiology, South Dakota State University, Brookings, SD 57007, USA

² Department of Agronomy, Horticulture and Plant Science, South Dakota State University, Brookings, SD 57007, USA

responses to drought stress could also have implication on selection for genotypes and development of agronomic practices to manage the environmental stress.

Much progress in understanding plant response to drought stress has been made in the model plant *Arabidopsis* with numerous molecular events, genes, and metabolites identified (Castro et al. 2012; Iordachescu and Imai 2008; Nakashima et al. 2014; Shinozaki and Yamaguchi-Shinozaki 2007). In contrast, knowledge of response to drought is limited and fragmented in crops such as soybean (*Glycine max* L. Merrill). Soybean is one of the most important crops worldwide for both protein meal and vegetable oil and accounts for approximately 58% the global oilseed production (Tran and Nguyen 2009). Drought is one of the major constraints limiting soybean production throughout the world and causes a significant loss of the soybean yield (Jin et al. 2018). A case study in Kentucky showed that up to 56 to 77% of the revenue from soybeans was lost due to the drought of recent years (Craft et al. 2015). In another aspect, timely and proper management decisions in response to drought stresses will allow soybean producers to mitigate yield losses. But drought damage on soybean is less visually dramatic as compared to other crops, such as maize. Furthermore, many cultivars are cultivated across the regions and years, and each of them can be different in response to drought. Therefore, it would be desirable to develop biomarkers for early diagnosis of drought stress. Recent transcriptomic studies on soybean response to drought stress, using microarray hybridization (Le et al. 2012; Tripathi et al. 2016) and RNA-Seq (Chen et al. 2016; Shin et al. 2015; Song et al. 2016), identified potential pathways associated with drought stress. Transcription factors (TFs) belonging to ERF, WRKYs, MYB, bHLH, and NAC families as well as plant hormonal pathways regulating the biosynthesis and sensitivity to abscisic acid (ABA) and jasmonate (JA) were identified to be key targets. Before soybean transcriptomic data were available, drought biomarkers were developed based on *Arabidopsis* gene expression data (Neves-Borges et al. 2012). The suitability of these markers in diagnostics of drought damage has not evaluated across a range of soybean genotypes. Thus, there is a need to develop sensitive and robust biomarkers for early diagnosis of drought stress based on soybean transcriptomic data and validate them in cultivars differing in drought tolerance.

This research was aimed at understanding early transcriptomic responses of soybean seedlings to drought stress under controlled conditions and developing biomarkers for early diagnostic of genotypic difference in drought sensitivity. Using microarray analysis, we identified ~700 drought-responsive genes functioning in 23 pathways. Subsequently, we developed drought response biomarkers by validating 10

DEGs from 5 genetic pathways in 20 commercial cultivars showing different levels of tolerance to drought. Thus, our analysis reveals an overall understanding of the gene expression and develops a new tool for decision making during drought stress in soybean. Here, we report the results and implications for improvement for soybean drought tolerance.

Materials and Methods

Plant Materials and Stress Treatments

A total of 21 soybean cultivars were selected for this research (Supplementary Table 1). The cultivar Williams 82 (W82), the genotype of the soybean reference genome sequence (Schmutz et al. 2010), was used for microarray analysis. W82 seeds were germinated, and seedlings were grown in pots filled with a mixture of vermiculite and perlite (Hummert International, San Louis, MO) at a ratio of 2:1 and kept in a growth chamber at 28 °C with a photoperiod of 16-h light and 8-h dark. Plants were watered with half-strength Hoagland solution every day to keep the medium moist. At Vegetative 3 (V3) stage when the third trifoliolate leaf has unrolled, seedlings were carefully removed from the media, cleaned with tap water, and gently wiped with tissue paper to remove the adhering water to the roots. The seedlings were placed on a stack of paper towels for 6 h to expose them to dehydration stress or placed in a beaker with distilled water for 6 h as controls. After the treatment, both the dehydration-treated and control seedlings were collected and frozen in liquid nitrogen for RNA extraction. Two biological replicates from seedlings grown under identical conditions at different times were collected, and three seedlings per sample were used.

The remaining 20 cultivars were used to validate drought-responsive biomarkers. Seedlings were planted in a 6-l pot containing soil media of vermiculite, perlite, and Sunshine Mix 1 (Sun Gro Horticulture, Agawam, MA) mixed at a 1:1:1 ratio and kept in a greenhouse room (continuous 28 °C temperature, photoperiod of 12/12 h day/night, 60% relative humidity) at the South Dakota State University (Brookings, South Dakota). Eight seeds were sown per pot and then thinned to four plants per pot at the V2 stage when the second trifoliolate leaf has unrolled. All pots were kept well-watered until the V3 stage. At this stage, drought stress was imposed for treated plants (V3-D) by withdrawing water for 14 days while all the control plants (V3-C) were well watered. At the end of the treatment, visual symptoms were scored and pictured, and plant heights were measured, and the youngest leaf was harvested from each plant for RNA isolation.

RNA Extraction, cDNA Synthesis, and Microarray Analysis

The frozen tissue was ground into fine powders in liquid nitrogen by mortar and pestle, and total RNA was extracted from the tissue powders using Trizol Reagent (Thermo Fisher Scientific, Waltham, MA) following the manufacturers' instructions. GeneChip Soybean Genome Array (Thermo Fisher Scientific), which contains 37,744 probe sets covering the entire soybean genome, was used for the hybridization. Preparation of complementary DNA(cDNA), dye labeling, microarray hybridization, and data acquisition were performed at the Integrated Genomics Facility of Kansas State University (Manhattan, KS) on a fee-for service basis.

For drought-responsive biomarkers, ~ 100 mg of the youngest leaf tissue was cut into small pieces, immediately transferred into a 2-ml tube containing 1 ml of TRI Reagent (Zymo Research, Irvine, CA), and ground with a metal bead using a TissueLyser II (Qiagen, Valencia, CA) at room temperature. Total RNA isolated from tissue-TRI Reagent mixture was purified by using a Direct-zol RNA MiniPrep Plus kit (Zymo Research), and the genomic DNA contamination was removed using on-column RNase-Free DNase I digestion following the manufacturer's protocols. RNA integrity was analyzed by electrophoresis on 1% agarose gels stained with ethidium bromide. RNA quantity was examined using a NanoDrop-1000 Spectrophotometer (Thermo Fisher Scientific) to ensure that A260/A280 ratios were ranged from 1.8 to 2.2 and A260/A230 ratios were above 2.0. Approximately 2 µg total RNA was used to synthesize the first-strand cDNA using M-MuLV Reverse Transcriptase (New England Biolabs, Ipswich, MA).

Data Visualization, Gene Ontology Annotation, and Pathway Analysis

The microarray datasets were analyzed using the DNA-CHIP (DCHIP) analyzer program, software for analysis and visualization of gene expression and SNP microarray data using model-based normalization and probe summarization methods (Li 2008; Li and Wong 2001). Genes differentially expressed in response to dehydration stress were selected with a minimum of twofold change in transcription, and a *P* value of < 0.05, and multiple testing was done using the permutation test provided within the program (100 times). The GeneChip Soybean Genome Array probe identifier was converted to Glyma1 model ID using the annotation tool at Soybase (<http://soybase.org/AffyChip/>). Annotations of the DEGs were obtained from Phytozome v11 (<https://phytozome.jgi.doe.gov/>) and National Center for Biotechnology Information (NCBI; <http://www.ncbi.nlm.nih.gov/>). Web Gene Ontology (GO) Annotation Plot (WEGO, <http://wego.genomics.org.cn/cgi-bin/wego/index.pl>) (Ye et al. 2006) was

used for plotting GO annotation results. Glyma1 ID and Glyma2 ID were inter-converted using the online identification converter tools (<http://www.soybase.org/correspondence/>). GO enrichment analysis of up- and down-regulated genes by using the AgriGO tool (<http://bioinfo.cau.edu.cn/agriGO/>) (Du et al. 2010). Pathway analysis was performed using MAPMAN software (Thimm et al. 2004) and DAVID analysis tools (<http://david.abcc.ncifcrf.gov/home.jsp>); (Huang et al. 2009; Huang et al. 2008). KOBAS 3.0 (http://kobas.cbi.pku.edu.cn/anno_iden.php) was used for KEGG pathway enrichment analysis (Xie et al. 2011).

Analysis of Cis-Acting Regulatory Elements and Prediction of Protein-Protein Interaction

Sequences of 1 kb upstream from the translation start site were retrieved as putative promoter region of the DEGs and used to search for transcriptional cis-elements using the PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). Abiotic stress response-related elements were chosen for expression analysis. For determining the potential functional relevance of the DEGs that are putatively involved in drought stress response, protein-protein interactions were predicted and analyzed using STRING software (<http://www.string-db.org/>).

Primer Design and Real-Time Quantitative PCR

Primers of quantitative real-time (reverse transcription) PCR (qRT-PCR) was designed using the Sigma-Aldrich OligoArchitect™ Online tool (<http://www.oligoarchitect.com>) to obtain a PCR product size of between 75 and 260 bp. Sequences of resultant primers (Supplementary Table 3) were used as queries for BLAST searches against the soybean genome at Gramene (<http://ensembl.gramene.org/Tools/Blast?db=core>) to verify their specificity.

qPCR was performed on a 7900HT Fast Real-Time PCR System (Thermo Fisher Scientific) using the following cycling parameters: 50 °C for 2 min, 95 °C for 10 min, and 40 cycles at 95 °C for 2 min, 60 °C for 30 s, and 72 °C for 30 s. Each amplification reaction contained 1 µL of cDNA from serial dilutions, 200 nM each forward and reverse primers, 0.4 µL 1× ROX Reference Dye LSR (passive reference), 10 µL of 2× SYBR Advantage qPCR Premix (Takara Bio USA, Mountain View, CA), and DNase-free water to a final volume of 20 µL. Data were collected during the extension phase, and dissociation curve analysis was performed by heating each reaction product from 60 to 95 °C and increasing the temperature with 0.5 °C increment over 5 s to verify the specificity of the primers. A negative control sample without an RNA template was also assayed to ensure that the samples were not contaminated. The soybean actin gene, *GmActin11* (*Glyma18g52780*), was used as a reference. Sample cycle

threshold (C_t) values were standardized for each template using the $2^{-\Delta\Delta C_t}$ method (Schmittgen and Livak 2008) was used to analyze relative changes in gene expression. Four biological replicates and two technical replicates were used to ensure statistical credibility.

Data Analysis and Statistics

A built-in statistics package was used for microarray analysis. The average and standard deviation (SD) for measured values from qRT-PCR and plant height were estimated using Microsoft® Excel functions. Student's t tests were performed using pooled SDs to evaluate the significance of the differences between the controls and the treatments. The software of Statistical Product and Service Solutions (SPSS) was used to calculate the correlation coefficient between plant height reduction percentage and the value of \log_2 fold change for biomarker development. The cut-off for statistical significance was set to a P value of 0.05 or less.

Results

Soybean Transcriptome Response to Dehydration

To obtain an overview of the early transcriptome changes in dehydration-stressed plants, we performed microarray analysis of dehydrated and control plants. Using three criteria, i.e., minimum of twofold change in transcription, $FDR \leq 0.05$, and P value < 0.05 , we identified a total of 697 differentially expressed genes (DEGs; 420 upregulated and 277 downregulated) in dehydration-stressed seedlings (V3-D) compared to control seedlings (V3-C) (Fig. 1a; Supplementary Table 2). Ten DEGs were selected for validation of the microarray data

by qRT-PCR using independent biological replicates. These 10 DEGs include eight upregulated genes and two downregulated genes functioning in various genetic pathways (Fig. 1b; Supplementary Table 3). Despite the slight difference, the expression profiles of these ten genes from the qRT-PCR experiment highly correlated with their expression data obtained from the microarray experiment ($r = 0.911$; Fig. 1c), indicating that the DEGs identified in the microarray hybridization experiment are reliable.

Functional Annotation of Drought-Responsive Transcripts

We conducted a GO analysis to annotate the genes and to identify functionally related genes that were differentially expressed in dehydrated soybean plants. Many DEGs were associated with response to stimulus processes and transcriptional regulatory functions, and there was significant variation between upregulated and downregulated genes related to these categories. Notably, genes associated with antioxidant properties were downregulated (Fig. 2; Supplementary Table 4).

We subsequently performed a GO enrichment analysis with ArgiGO. The 420 upregulated genes formed 22 functional groups (Supplementary Figs. 1A and 1B; Supplementary Table 5). Three of these 22 groups, i.e., response to stress (GO: 0006950; 36 genes), response to water deprivation (GO: 0009414; 11 genes), and response to ABA stimulus process (GO: 0009737; 13 genes), were widespread in the biological process category. In the molecular function category, we detected one prominent functional gene group involved in phosphoprotein phosphatase activity (GO: 0004721; 11 genes) (Supplementary Fig. 1B; Supplementary Table 5). Of the 277 downregulated genes, one prominent functional gene group was detected in the biological process category, which

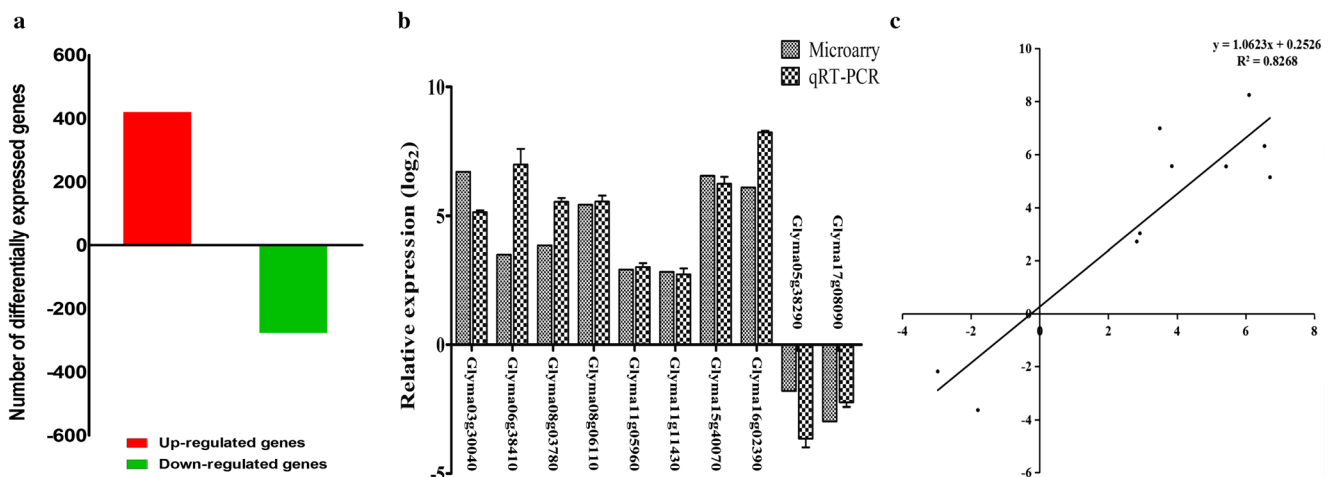
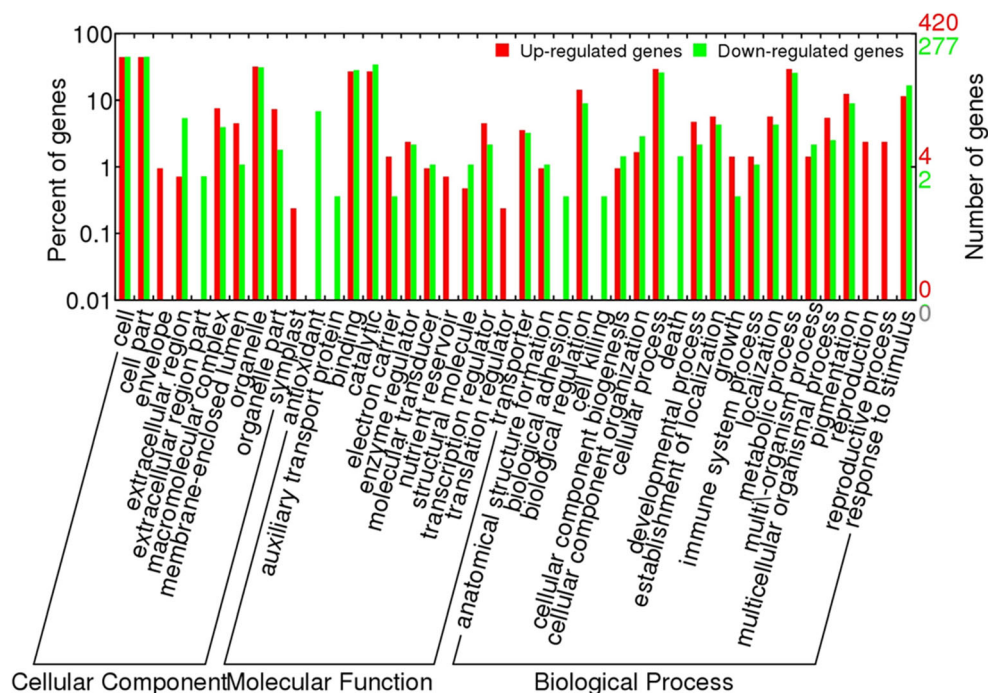


Fig. 1 Analysis and validation of microarray data. **a** Total number of DEGs, transcript abundance of which significantly decreased or increased under dehydration (2-fold at P value < 0.05). **b** qRT-PCR

verification of 10 DEGs. The gene loci are indicated on the x axis, and their \log_2 fold changes are indicated on the y axis. **c** Correlation between microarray (x-axis) and qRT-PCR data (y-axis) for the ten selected DEGs

Fig. 2 GO analysis of DEGs in W82 seedlings under dehydration. GO categories under the cellular component, molecular function, and biological process are indicated on the x axis; gene category frequencies for each domain are indicated on the left y axis; and the number of genes in each category on the right y axis



was associated with response to oxidative stress (GO: 0006979 and consisted of 21 genes (Supplementary Fig. 2A; Supplementary Table 5). In the molecular function category, one prominent functional gene class was found associated with peroxidase activity (GO: 0004601), and 19 genes were contained in this category (Supplementary Fig. 2B; Supplementary Table 5). In the cellular component category, we identified two prominent functional gene classes, namely extracellular region (GO: 0005576; 15 genes) and plant-type cell wall (GO: 0009505; six genes) (Supplementary Fig. 2C; Supplementary Table 5). This result is consistent with the downregulation of antioxidant-associated genes (Fig. 2; Supplementary Table 4). These results suggest that soybean mobilize stress response pathways via ABA and reactive oxygen species (ROS) signals.

Using the David data tool (National Institute of Allergy and Infectious Diseases, NIH), we categorized the 277 downregulated genes into seven major groups, the similarity with GO enrichment analyses. Among these groups of genes, group 1 consisted of peroxidase-related genes, which is consistent with the GO enrichment data (Supplementary Table 5). In addition, group 4 and group 7 were comprised of genes related to protein kinase and sugar transporter. In group 6, genes were mainly associated with AP2/ERF TFs (Supplementary Table 6). Similarly, 422 upregulated genes were categorized into nine major groups. Among these groups of genes, group 1 consisted of protein phosphatase 2C genes, which is consistent with the GO enrichment data (Supplementary Table 5). Group 3 was comprised of genes encoding TFs. In group 9, genes were mainly associated with protein kinase (Supplementary Fig. 4; Supplementary Table 6).

We evaluated the DEGs in key metabolic and signaling pathways in a systematic manner using MAPMAN and KEGG (Supplementary Fig. 3). The result showed that the upregulated genes were grouped into 24 categories and downregulated genes classified into 22 categories (Fig. 3a). Of these categories, the “transport” and “transcription factors” were the two most abundant categories in both up- and downregulated genes. Twenty-six genes were grouped into the “hormone metabolism” category.

Dehydration-responsive genes were significantly enriched in 23 KEGG pathways (Fig. 3b). The “metabolic pathways,” “biosynthesis of secondary metabolites,” “plant hormone signal transduction” and “plant-pathogen interaction” are the major enrichment pathways. The metabolic pathways, which contains the largest number of DEGs, includes biosynthesis of secondary metabolites, “carbohydrate metabolism,” “lipid metabolism,” and “amino acid metabolism.” In the “signal transduction” category, plant hormone signal transduction and “phosphatidylinositol signaling system.” In the hormone-signaling transduction pathway, ABA-signaling pathway was enriched. These results revealed that the DEGs mainly involved in metabolic pathways and plant hormone signal transduction, which primarily encoded TFs, protein kinases, hormone signaling, and other regulatory proteins. These DEGs will be addressed in details as follows.

Transcription Factors

A total of 60 DEGs were annotated as TF-coding genes, 41 of which were upregulated and 19 downregulated by dehydration (Supplementary Fig. 4; Supplementary Table 7). These 60

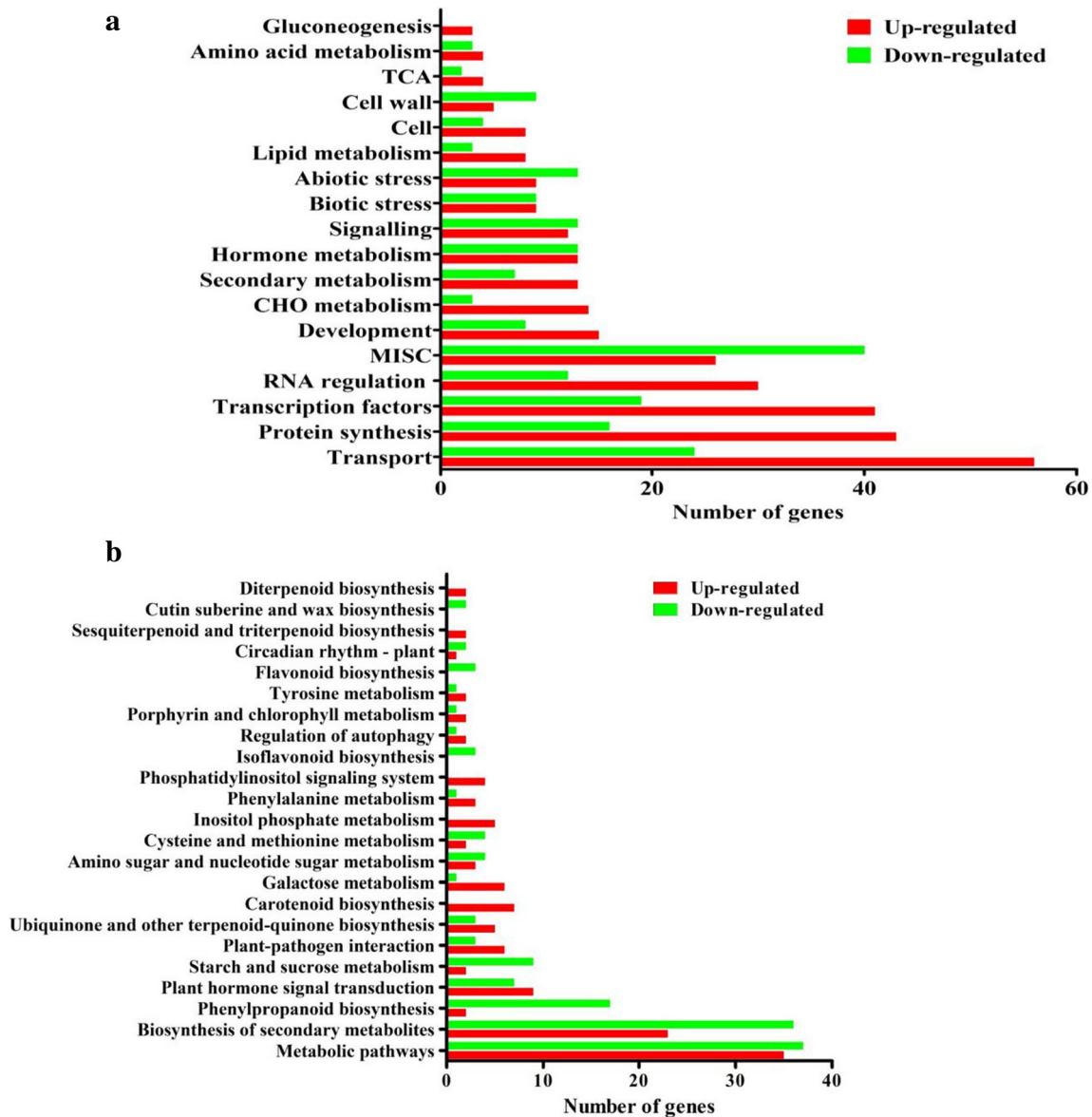


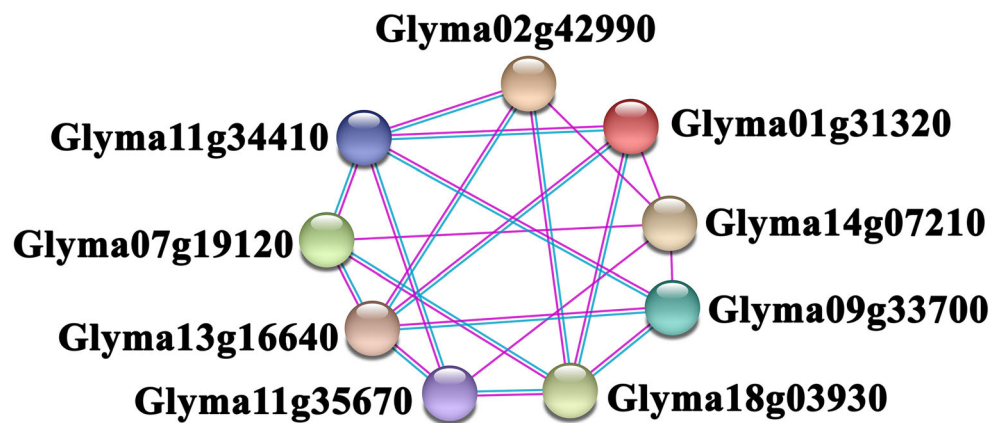
Fig. 3 Distribution of DEGs into major biological processes and KEGG pathway. **a** Distribution of DEGs into major biological processes classified using MapMan. The number of genes in each biological process is indicated on the x axis, and categories of biological

processes are indicated on the y axis. **b** KEGG pathway enrichment of DEGs. The number of genes in each KEGG pathway is indicated on the x axis, and the categories of KEGG pathway are indicated on the y axis

TF DEGs belong to seven major groups, i.e., AP2/ERF, bHLH, bZIP, HD-Zip, MYB, NAC, and zinc finger families. Group 1 contained 12 auxin response factor/ethylene-responsive factor (ARF/ERF) genes. Six of these genes were induced, and the other six were suppressed under dehydration. *Glyma14g06080*, soybean homolog of *AtDREB2C*, and *Glyma12g33020*, soybean homolog of *AtDREB5*, were induced 3.85- and 6.51-fold under dehydration stress, respectively. Group 2 consists of nine basic helix-loop-helix (bHLH), and three of them were upregulated. Group 3 is comprised of three upregulated genes encoding basic zipper (bZIP) transcription factors, of which *Glyma06g04350*, soybean homolog of *AtABI5*, was upregulated 7.60-fold. Group 4

included four HD-Zip genes, three of which were upregulated and one of which was downregulated by dehydration. Group 5 was composed of MYB family genes, of which the transcription level of three genes was increased, and three MYB-related genes were decreased by dehydration. Group 6 consists of 11 NAC family genes, including one soybean homolog of *NAC072/RD26* gene. Most notably, all of these 11 NAC TF genes were upregulated in dehydrated W82 seedlings. Finally, group 7 consists of four zinc finger family genes, which include two C2H2 genes and two C3H genes. The remaining 11 TF genes encoded two MADS_MIKC, two WRKY, one NF-YC, one ARR-B, one LBD, one heat shock factor (HSF), one KNOX, one PLATZ, and one ZF-HD gene.

Fig. 4 Predicted protein-protein interaction network of DEGs. Protein-protein interaction network of PP2Cs in upregulated genes interact with ABA receptors in downregulated genes. The key is a list at the right of the figure



Alteration of Hormonal Signals

Phytohormones, including ABA, auxin (IAA), brassinosteroids (BRs), cytokinin (CK), ethylene (ET), gibberellin (GA), and JA, virtually regulate every process of plant growth and development. MAPMAN and KEGG pathway analysis showed that all these phytohormones were involved in dehydration response with 43 DEGs identified, of which 22 are related to ABA, 9 to GA, and 6 to ET pathways (Fig. 3; Supplementary Fig. 4). Of the 22 ABA-related genes, 4 are involved in ABA synthesis, 11 encode protein phosphatase 2Cs (PP2Cs), 2 encode ABA-responsive element-binding factors (ABFs), and 5 encode ABA receptors (PYL/PYR). Interestingly, while all the four ABA biosynthetic genes, 11 PP2C, and 2 ABF genes were induced or upregulated by dehydration, the 5 PYL/PYR genes were downregulated. Of the six ET-related genes, four genes were annotated as a 1-aminocyclopropane-1-carboxylate oxidase, and one gene encodes 1-aminocyclopropane-1-carboxylate synthase, all of which catalyzes ET biosynthesis. Nine DEGs were involved in GA synthesis and signal pathway, four upregulated and five downregulated.

Protein Kinases

A total of 37 DEGs encoding protein kinases were identified in dehydrated W82 seedlings (Supplementary Fig. 4; Supplementary Table 6). Of these 37 protein kinases, 28 are receptor-like kinases-Pelle, 12 genes upregulated and 16 downregulated. The second prominent group of protein kinases is composed of eight calmodulin-dependent protein kinases with three upregulated and five downregulated by dehydration. Finally, *Glyma08g23900* encodes a mitogen-activated protein kinase kinase 5-like, which belongs to STE group, and was induced by dehydration.

Transporter Proteins

Dehydration induced the expression of transport-related proteins such as ATP-binding cassette (ABC) transporters, sugar transporters, water channels (Supplementary Fig. 4). Five genes encoding ABC transporter proteins and two genes encoding aquaporin proteins were upregulated, and one gene was downregulated. We also identified five upregulated and six downregulated sugar transporters under the dehydration stress.

Induction of Chaperone Proteins

Proteins like late embryogenesis-abundant (LEA) proteins and heat shock proteins (HSPs) are accumulated during stress and function in protective mechanisms. In this study, nine DEGs encode LEA proteins with five upregulated and four downregulated by dehydration (Supplementary Fig. 4). Three transcripts coding for an HSP20, an HSP70, and a DnaJ20 were found induced under dehydration, but an HSP21-coding gene was downregulated (Supplementary Fig. 4).

Oxidases under Dehydration Stress

Oxidative stress arises as a result of ROS accumulation during the dehydration stress. Plant cells are protected from the oxidative damage during stress conditions by different enzymatic antioxidant mechanisms. These enzymes include CYP450s, glutathione S transferases (GSTs), peroxidases, and esterases (Supplementary Fig. 4). We found that genes encoding 11 cytochrome P450s and two glutathione S-transferases were differentially expressed under dehydration stress (Supplementary Table 7). While all five lipases were upregulated under dehydration, all 16 peroxidase genes and 3 GDSDL-motif lipase genes were downregulated.

Carbohydrate Metabolism

KEGG analysis revealed that the terms of “galactose metabolism,” “starch and sucrose metabolism,” and “inositol phosphate metabolism” were enriched with 23 DEGs identified in carbohydrate metabolism pathway (Supplementary Fig. 4). Expression of three genes for galactose degradation, three genes for pectin degradation, and three genes for cell wall degradation was reduced, and transcription of two cell wall formation genes was increased. Genes for the synthesis of stachyose, raffinose, and trehalose and degradation of starch were upregulated, but six genes for sucrose metabolism were downregulated. In the “inositol phosphate metabolism” category, five myo inositol-3-phosphate synthesis-related genes, three for myo inositol-3-phosphate synthase (MIPS), and two for phosphoinositide phospholipase C (PLC) were induced by dehydration. These results suggest reinforcement of cell wall, increase of soluble sugars, and possibly active lipid metabolism occurred during dehydration stress.

Predicted Protein-Protein Interaction

GO enrichment, MAPMAN, and KEGG pathway analysis showed that ABA-signaling pathway is specific in the biological process of dehydration response (Fig. 3; Supplementary Table 5). These dehydration-responsive genes were predicted to undergo protein-protein interactions using STRING. We found that seven upregulated genes encoding protein phosphatase 2C can interact with five downregulated ABA receptors PYLs (Fig. 4). All of these interactions suggest that dehydration-responsive phosphatase 2C play a significant role in the ABA-mediated signaling pathway.

Cis-Regulatory Elements in the Promoters of Transcription Factors

To investigate whether certain promoter motifs of TF genes were involved in the different pathways to respond to drought stress, we surveyed the sequence 1000 bp upstream of each TF gene for *cis*-motifs. Fourteen plant hormone-related and drought stress response-related *cis*-motifs were detected in the gene-promoter sequences of 39 upregulated and 19 downregulated TF genes (Supplementary Table 8). The major *cis*-regulatory elements identified are ABRE (ABA-responsive element), GARE (gibberellin-responsive element), MBS (MYB binding site involved in drought-inducibility), TC-rich (*cis*-acting element involved in defense and stress responsiveness), and TCA elements (*cis*-acting element involved in salicylic acid responsiveness). Interestingly, ABRE regulatory element was identified in the promoter regions of 28 upregulated and three downregulated TF genes and accounting for 71.79 and 15.79% of the total TF genes, respectively.

Development of Drought-Responsive Biomarkers

We screened 20 soybean cultivars or breeding lines for the difference in response to drought stress to evaluate the suitability of selected genes as biomarkers. After 14 days of water withdrawal, these cultivars showed a significant difference in leaf wilting (Fig. 5a) and plant height reduction (Fig. 5b) as compared to their well-watered control plants (Supplementary Table 1). Based on the degree of leaf wilting, 20 soybean cultivars were divided into four groups (Supplementary Table 1). Under drought stress, no obvious leaf wilting was detected in drought-stressed plants of the group-1 cultivars (A_11, Triall, 91Y74R2Y, and Brookings). Only slight wilting and dark green was observed in leaves of the drought-treated group-2 cultivars (SD2172R2Y, Stine0480, S12-L5, S14-J7, S17-B3 and S18-C2). A moderate leaf wilting was of in the group-3 cultivars (Roberts, Surge, S06-H5, S08-61, S10-G7, and Codington). Severe leaf wilting was found in the group-4 cultivars (Deuel, SD1093RR, SD2091R2Y and S10-P9). In addition to the variation in the leaf wilting, significant differences were also observed in plant height between control and drought-stressed plants. The degree of plant height reduction and degree of leaf wilting showed a low level of correlation ($r = 0.406$), suggesting different yet overlapping mechanisms might underlie variations in these traits. Thus, the 20 cultivars were grouped based on a percentage of plant height reduction and *P* values (Supplementary Table 1). Considering both leaf wilting and plant height reduction, we selected two drought-tolerant genotypes, Triall and S12-L5, and two drought-sensitive genotypes, SD1093RR and Deuel, for validating drought-sensitivity biomarkers because they showed consensus performance in both traits.

Because drought stress induced the high expression of genes encoding dehydrins, LEA proteins, TFs, protein kinases, and other regulatory proteins, we selected 10 DEGs from these categories to determine drought sensitivity (Supplementary Table 2). RT-qPCR results indicated that these ten markers effectively confessed the genotypic difference in drought-induced expression level between drought-tolerant genotype and drought-sensitive genotype (Fig. 5c). Eight of these ten genes were upregulated in our microarray data, i.e., *Glyma03g30040*, *Glyma04g01181*, *Glyma08g06110*, *Glyma11g05960*, *Glyma11g11430*, *Glyma12g22880*, *Glyma15g40070*, and *Glyma16g02390* showed the most significant difference between V3-D and V3-C plants and exhibited consistent differences between drought-tolerant and drought-sensitive genotypes. The drought-induced expression of these eight genes is much higher in the drought-sensitive cultivars SD1093RR and Deuel as compared to the drought-tolerance cultivars Triall and S12-L5, from 2.8-fold in gene *Glyma11g05960* encoding GmZEP to 126.6-fold in gene *Glyma04g01181* encoding a dehydrin. The remaining two genes, *Glyma08g03780* and

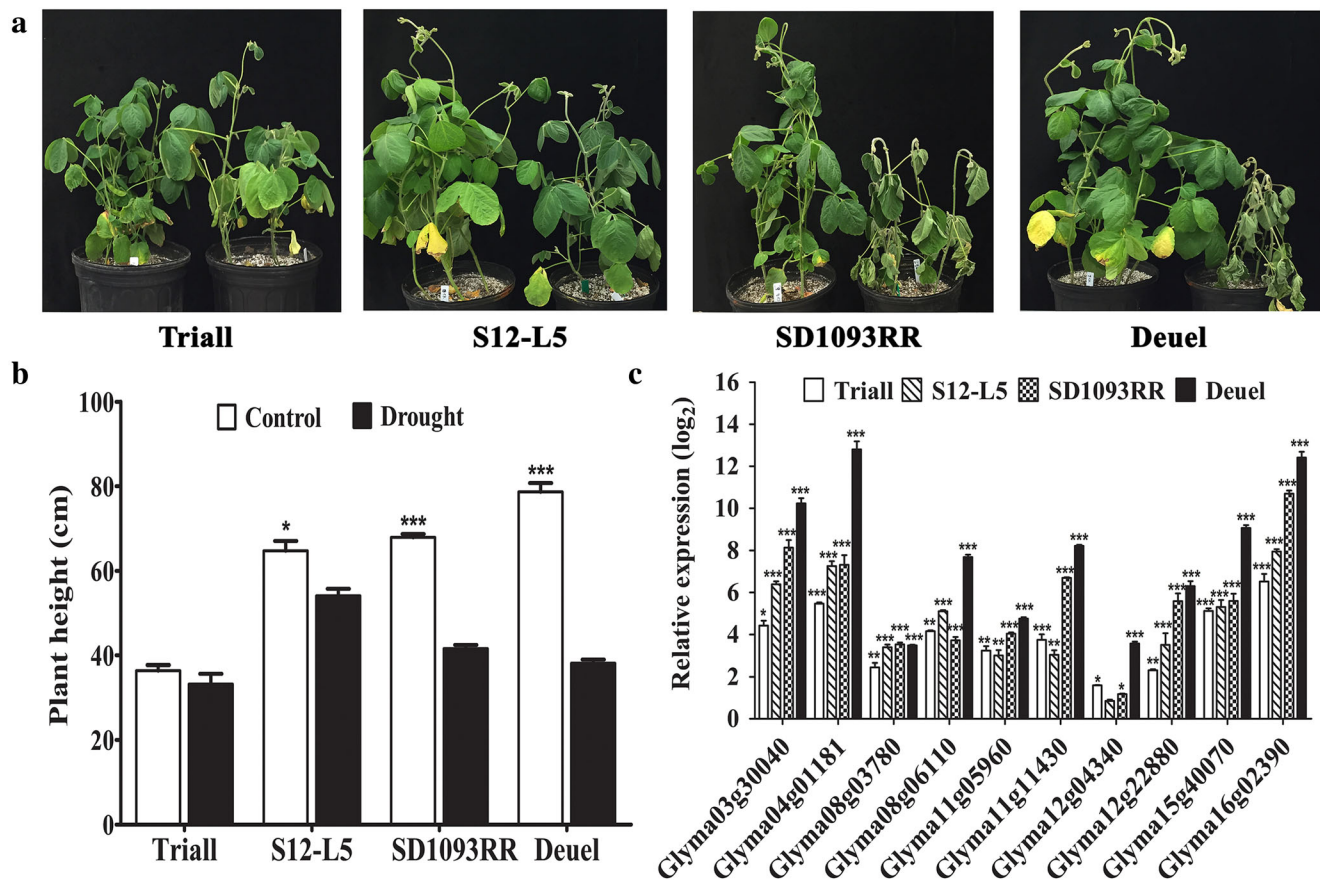


Fig. 5 Development of drought-responsive biomarkers. **a** Phenotypes of two genotypes under well-watered (left) and drought-stressed (right) conditions. **b** Plant height of two genotypes under well-watered (yellow) and drought stress (blue) conditions. Names of the soybean cultivars tested are indicated on the x axis, and their plant heights (cm)

are indicated on the y axis. **c** RT-PCR assay of 10 selected genes in the drought-resistant and drought-sensitive cultivars. The gene loci are indicated on the x axis, and log₂ fold changes relative to the control plants are indicated on the y axis. The *GmActin11* was used as a reference. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Glyma12g04340, showed an obviously different expression pattern as compared to other eight genes among four soybean cultivars. Five of the eight genes, i.e., *Glyma03g30040*, *Glyma11g05960*, *Glyma11g11430*, *Glyma12g22880*, and *Glyma16g02390*, showed a very high correlation between the expression data and plant height reduction ($r > 0.953$, $P < 0.047$), indicating their suitability as biomarkers to determine drought sensitivity and tolerance.

Discussion

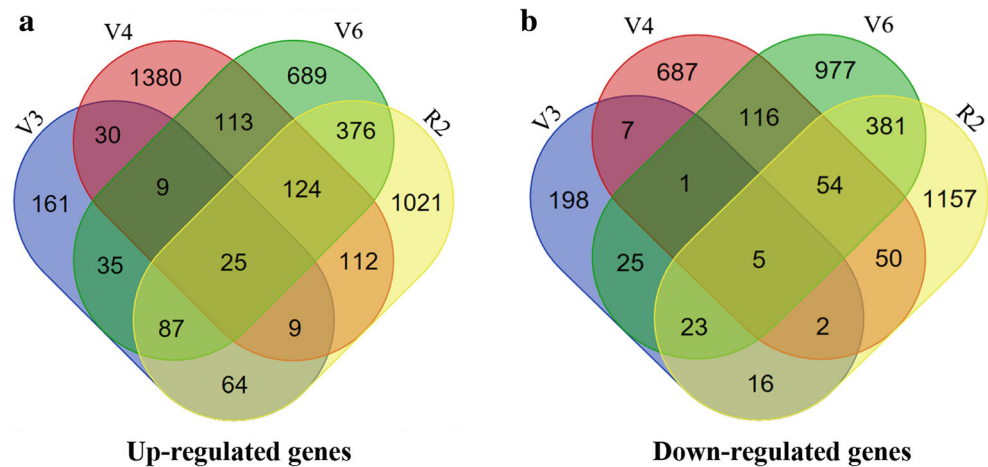
Dehydration adversely affects many biological processes of plant life and trigger multiple genetic pathways. Thus, tolerance to drought, prolonged slow dehydration, is one of the most important yet most complicated traits for crops like soybean. In this study, our microarray analysis identified ~700 candidate genes responsive to dehydration stress in 14-day-old seedlings of the soybean plants. Our research highlights TFs and ABA signaling in

soybean plant response to dehydration treatment. From the candidate gene pool, we selected five genes as biomarkers to predict plant responses to drought stress. All this is of theoretic and applied significance for improving drought tolerance and managing drought damages in crops.

Effect of Transcriptome-Profiling Platforms and Treatment Methods for DEG Identification

Microarray and RNA-Seq are powerful platforms to reveal transcriptome response to abiotic stress such as drought. Recently, drought-responsive genes have been identified in soybean W82 leaf tissues under different drought conditions by using the Affymetrix microarray at V6 and R2 stages (Le et al. 2012) and RNA-Seq platform at V4 stage (Chen et al. 2016). In the present study, DEGs were identified from soybean W82 seedlings at V3 stage under 6-h dehydration. A Venn diagram analysis was performed to compare the DEGs at the V3 stage and those from other developmental stages.

Fig. 6 Venn diagram of DEGs in the different development stage of soybean W82 under drought stress. Venn diagram of upregulated genes **a** and downregulated genes **b** in different development stages under different drought condition



The result showed that overlap between V3 and V4, between V3 and V6, and between V3 and R2 stages is 17.38%, 37.14%, and 43.81%, respectively, for the upregulated DEGs (Fig. 6a) and 5.42%, 19.49%, and 16.61%, respectively, for the downregulated DEGs (Fig. 6b). The percentages of overlap genes between V3 and V4 are expected highest because the physiological condition between these two stages is very close. The opposite, however, holds true, and overlap in both up- and downregulated genes were lowest among three comparisons. The DEGs from V3-D/V3-C, V6-D/V6-C, and R2-D/R2-C comparisons were identified using the microarray platform, but the DEGs from comparison V4-D/V4-C were identified by using RNA-Seq. Thus, the low DEG overlap between the V3 and V4 could also be due to different methods used for drought treatment and transcriptome evaluation. Drought stress was imposed by withdrawing water for 7 days in the V4, V6, and R2 stages, while soybean plants were dehydrated for 6 h in the V3 stage in the present study. Validation of 10 DEGs induced by dehydration using qRT-PCR in drought-stressed plants confirmed the DEGs identified are also induced by drought. Although they were identified in a relatively short time treatment, the expression patterns of these DEGs well corresponded to the prolonged slow dehydration process (drought) based on the qPCR assays in the four selected varieties.

Transcription Factors Responding to Drought Stress

We identified a total of 60 dehydration-regulated TFs of seven major families (Supplementary Table 6), two-thirds of which were upregulated (Fig. 3; Supplementary Table 2). Of these TF DEGs, AP2/ERF, bHLH, HD-Zip I, and NAC families are the most significant. All the 11 NAC TF DEGs were upregulated by dehydration probably because NAC TFs regulate the development of water-conducting xylem tissues (Xu et al. 2014; Zhong et al. 2006; Zhong et al. 2007). Among these NAC TFs, *Glyma12g22880*, a homolog to *RD26* of

Arabidopsis (Le et al. 2012), was most induced, 13.2-fold 6 h after dehydration and ~80-fold after 14 days of water withdrawal. The transcription levels of the remaining homologs of *AtRD26*, i.e., *Glyma06g38410*, *Glyma12g35000*, and *Glyma13g35550*, remained unchanged under dehydration stress, indicating that functional diversification among four *GmRD26* homologs occurred during soybean evolution. *AtRD26* (Fujita et al. 2004) and its ortholog genes in tea (Wang et al. 2016), potato (Singh et al. 2013), cucumber (Zhang et al. 2017), and canola (Ying et al. 2014) are highly responsive to abiotic stress. In *Arabidopsis*, RD26 acts as a transcriptional activator in ABA-dependent dehydration responses (Fujita et al. 2004), while it mediates suppressive crosstalk with BR signaling (Ye et al. 2017). Our results corroborate its function in soybean drought response. The most induced bHLH TF gene is *Glyma11g05810*, encoding a SPATULA-like (SPT) TF. SPT negatively regulates the growth of plant organs in *Arabidopsis* (Makkena and Lamb 2013). Increased transcription of *Glyma11g05810* may explain the effect of drought on reduced plant height (Fig. 5). TF *Glyma17g10290*, encoding a bHLH79-like protein, was found to be induced 2.8-fold in the seedlings of drought-sensitive cultivar W82 under dehydration stress, confirming its previously reported role in soybean drought response (Osorio et al. 2012). Of the six upregulated AP2/ERF TF genes, three *DREB* homologs of soybean, i.e., *GmDREB2* (*Glyma06g04490*), *GmDREB5-like* (*Glyma12g33020*), and *DREB2A* (*Glyma14g06080*), are most prominent, confirming their induction by drought stress in previous soybean studies (Chen et al. 2007; Marcolino-Gomes et al. 2015; Marcolino-Gomes et al. 2013; Maruyama et al. 2012; Mizoi et al. 2013). Within HD family, *GmHDZ72*, orthologous to *Arabidopsis* HD-Zip I gene *ATHB12*, was most highly induced by dehydration among all the TF DEGs. *ATHB12* play important roles in response to drought stress. It activates protein phosphatase 2C (PP2C) genes and represses ABA receptor genes (Valdés et al. 2012) as a negative regulator of ABA signaling, but it

also negatively regulates the plant growth by decreasing the expression of a gibberellin 20-oxidase gene (Olsson et al. 2004; Son et al. 2010). All these suggested that transcription regulation mechanisms in response to dehydration/drought stress are conserved in the land plants.

Hormonal Homeostasis and Signals in Response to Drought Stress

Phytohormone ABA is central in mediating drought stress signal. Homeostasis of endogenous ABA is maintained by ABA synthesis and catabolism. The zeaxanthin epoxidase (ZEP) (Marin et al. 1996) and 9-cis-epoxycarotenoid dioxygenase (NCED) (Schwartz et al. 1997; Seki et al. 2007) are two key enzymes in controlling ABA biosynthesis. The overexpression of *AtNCED3* (Iuchi et al. 2001) and *BnNCED3* in Arabidopsis (Xu and Cai 2017) and *CsNCED3* in tobacco (Pedrosa et al. 2017) enhanced drought tolerance. Our results showed that soybean NECD-coding gene *Glyma15g40070* was induced 93.48-fold after 6 h dehydration. Both ZEP-coding genes *Glyma09g00260* and *Glyma11g05960* were increased 3.01- and 7.56-fold, respectively, under dehydration stress (Supplementary Table 2). In another aspect, transcription of *Glyma01g35660* encoding ABA 8'-hydroxylase CYP707A1b, which inactivate ABA (Saito et al. 2004), was induced 4.5-fold in seedlings by dehydration. This result is consistent with the previous study, which reported that dehydration induced expression of soybean *CYP707A1b* (Zheng et al. 2012). Compared to the *NCED* gene, induction of the ABA catabolic gene is much lower, suggesting a net increase in the levels of bioactive ABA form by dehydration, similar to observation in other plants (Seki et al. 2007). Increased transcription of ABA biosynthetic genes, however, is not necessarily correlated with drought tolerance. For example, the expression level of *GmNECD* (*Glyma15g40070*) and *GmZEP* (*Glyma11g05960*) are higher in the drought-sensitive lines than in drought-tolerant lines (Fig. 5b). Elevated transcription of the ABA synthetic genes suggested that the drought-sensitive lines are low in ABA sensitivity and that reduced ABA sensitivity and increased drought sensitivity further fuels up the ABA biosynthesis in the drought-sensitive genotypes.

The core ABA-signaling pathway consists of three core components: ABA receptors pyrabactin resistance 1 (PYR1) and PYR1-like (PYL), PP2C, and SNF1-related protein kinase 2 (SnRK2), in which PYR1/PYLs and SnRK2s are positive regulators, and PP2Cs are negative regulators of the ABA signal (Zhu 2002, 2016). If the transcription of these genes directly contributes to their protein activities, transcription of *PYR1/PYLs* and *SnRK2s* is expected to increase and *PP2Cs* to decrease upon dehydration treatment. Different from this scenario, our pathway enrichment study showed that five ABA receptor genes were downregulated, 11 *PP2C* genes were upregulated, and no change in transcription level was

observed for *SnRK2* genes (Supplementary Table 6). This opposite patterns for PYLs and PP2Cs matches the ATHB12/PYL/PP2C module as part of negative feedback of the ABA-signaling pathway (Valdés et al. 2012) because *GmHDZ72*, the *ATHB12* ortholog gene of soybean, was highly induced by dehydration (Supplementary Table 6). Further support to this ATHB12/PYL/PP2C module came from protein-protein interaction prediction, which revealed that five downregulated GmPYLs interacted with 11 highly upregulated GmPP2Cs (Fig. 4a) and negatively regulated the classic ABA-PYL-PP2C-SnRK2 signaling pathway. Consistent with this prediction, the cis-element analysis found that *GmHDZ72* and 10 of 11 PP2C genes contain the ABRE motif in their promoters (Supplementary Table 8). All these results suggest that the negative feedback loop plays an important role in fine tuning ABA signal under drought stress.

Transcription of genes for ET and GA metabolism are also altered by the dehydration treatment but showed a bidirectional pattern (Supplementary Table 6). Expression of *Glyma02g15380*, coding for GA20ox for a rate-limiting step in GA biosynthesis, increased 4.3-fold, but transcription of GA catabolic genes *Glyma13g33290* and *Glyma15g10070*, both encoding 2- β -dioxygenase (GA2OX), increased 5- and 2.3-fold, respectively, suggesting a reduction of bioactive GA under dehydration stress. ET biosynthetic gene *Glyma08g02130*, coding for 1-aminocyclopropane-1-carboxylate synthase (ACS), and *Glyma07g25390*, encoding 1-aminocyclopropane-1-carboxylate oxidase (ACO), were upregulated ~4-fold, but three ACO-coding genes regulated in this study (Supplementary Table 6). The observation of upregulation of both biosynthesis and catabolic genes for ABA and bidirectional changes of the ET and GA biosynthesis gene expression may implicate spatial fine tuning on homeostasis of these hormones in response to dehydration stress.

Drought Sensitivity Biomarkers

Biomarkers are widely used in diagnostic of human diseases and plant biotic stress but rarely used in monitoring plant responses to abiotic stress. This research selected ten DEGs as candidates to develop diagnostic markers for plant early responses to drought stress. Opposite to our expectation, the drought induction level of these genes was much higher in the drought-sensitive genotypes, suggesting that they are drought sensitivity biomarkers. Five drought biomarkers, i.e., *GmLEA4-like* (*Glyma03g30040*), *GmZEP* (*Glyma11g05960*), *GmERD7-like* (*Glyma11g11430*), *GmRD26* (*Glyma12g22880*), and *GmHDZ72* (*Glyma16g02390*), showed similar expression pattern and highly correlated with drought-tolerant and drought-sensitive phenotypes (Figs. 5 and 6). The last four genes function in the ABA pathway, one coding for ABA synthesis enzyme, one for ABA-responsive protein, and two for TFs. These four genes

were upregulated to different degrees in drought-tolerant and drought-sensitive cultivars from 32.43- to 544.57-fold change under drought stress treatment (Fig. 5c). This observation is consistent with what was found in *Arabidopsis*, where drought stress stimulates ABA biosynthesis (Fujita et al. 2011). Compared to the drought tolerance marker, the drought sensitivity markers are more suitable for diagnostic of the drought damage. Different from previously reported drought marker genes in soybean roots (Neves-Borges et al. 2012), these five marker genes were assayed in soybean leaves. Thus, they can be more conveniently used for diagnostic of the degree of drought damage. Because these five marker genes were expressed at a much higher level in the sensitive cultivars as compared to the tolerant cultivars in response to drought stress, increase in their expression is expected to be detected in the early stages of drought stress. With the modification of sampling method, these drought-sensitivity markers can be used for on-farm diagnostics and for screening drought tolerance to reduce the experiment cycle time.

Funding Information This research was supported by South Dakota Excellence Center for Drought Research and South Dakota Soybean Research and Promotion Council. The use of instruments available at South Dakota State University's Functional Genomics Core Facility and BioSNTR supported in part by the National Science Foundation/EPSCoR grant no. 0091948 and Cooperative Agreement no. IIA-1355423, and by the State of South Dakota, are gratefully acknowledged.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

References

- Castro PH, Tavares RM, Bejarano ER, Azevedo H (2012) SUMO, a heavyweight player in plant abiotic stress responses. *Cell Mol Life Sci* 69:3269–3283
- Chai HH, Ho WK, Graham N, May S, Massawe F, Mayes S (2017) A cross-species gene expression marker-based genetic map and QTL analysis in bambara groundnut. *Genes (Basel)* pii:E84
- Chen M, Wang QY, Cheng XG, Xu ZS, Li LC, Ye XG, Xia LQ, Ma YZ (2007) GmDREB2, a soybean DRE-binding transcription factor, conferred drought and high-salt tolerance in transgenic plants. *Biochem Biophys Res Commun* 53:299–305
- Chen W, Yao Q, Patil GB, Agarwal G, Deshmukh RK, Lin L, Wang B, Wang Y, Prince SJ, Song L, Xu D, An YC, Valliyodan B, Varshney RK, Nguyen HT (2016) Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. *Front Plant Sci* 7:1044
- Collins NC, Tardieu F, Tuberosa R (2008) Quantitative trait loci and crop performance under abiotic stress: where do we stand? *Plant Physiol* 147:469–486
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* 339:819–823
- Craft KE, Mahmood R, King SA, Goodrich G, Yan J (2015) Twentieth century droughts and agriculture: examples from impacts on soybean production in Kentucky, USA. *Ambio* 44:557–568
- Du Z, Zhou X, Ling Y, Zhang Z, Su Z (2010) agriGO: a GO analysis toolkit for the agricultural community. *Nucleic Acids Res* 38(Web Server issue):W64–W70
- Fujita M, Fujita Y, Maruyama K, Seki M, Hiratsu K, Ohme-Takagi M, Tran LS, Yamaguchi-Shinozaki K, Shinozaki K (2004) A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. *Plant J* 39:863–876
- Fujita Y, Fujita M, Shinozaki K, Yamaguchi-Shinozaki K (2011) ABA-mediated transcriptional regulation in response to osmotic stress in plants. *J Plant Res* 124:509–525
- Huang DW, Sherman BT, Stephens R, Baseler MW, Lane HC, Lempicki RA (2008) DAVID gene ID conversion tool. *Bioinformatics* 24:428–430
- Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4:44–57
- Iordachescu M, Imai R (2008) Trehalose biosynthesis in response to abiotic stresses. *J Integr Plant Biol* 50:1223–1229
- Iuchi S, Kobayashi M, Taji T, Naramoto M, Seki M, Kato T, Tabata S, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K (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
- Jin Z, Ainsworth EA, Leakey ADB, Lobell DB (2018) Increasing drought and diminishing benefits of elevated carbon dioxide for soybean yields across the US Midwest. *Glob Chang Biol* 24:e522–e533
- Le DT, Nishiyama R, Watanabe Y, Tanaka M, Seki M, Le HH, Yamaguchi-Shinozaki K, Shinozaki K, Tran LS (2012) Differential gene expression in soybean leaf tissues at late developmental stages under drought stress revealed by genome-wide transcriptome analysis. *PLoS One* 7:e49522
- Li C (2008) Automating dChip: toward reproducible sharing of microarray data analysis. *BMC Bioinformatics*. 9:231
- Li C, Wong W (2001) Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci USA* 98:31–36.
- Makkena S, Lamb RS (2013) The bHLH transcription factor SPATULA is a key regulator of organ size in *Arabidopsis thaliana*. *Plant Signal Behav* 8:e24140
- Marcolino-Gomes J, Rodrigues FA, Oliveira MC, Farias JR, Neumaier N, Abdelnoor RV, Marcelino-Guimarães FC, Nepomuceno AL (2013) Expression patterns of GmAP2/EREB-like transcription factors involved in soybean responses to water deficit. *PLoS One* 8:e62294
- Marcolino-Gomes J, Rodrigues FA, Fuganti-Pagliarini R, Nakayama TJ, Ribeiro Reis R, Bouças Farias JR, Harmon FG, Correa Molinari HB, Correa Molinari MD, Nepomuceno A (2015) Transcriptome-wide identification of reference genes for expression analysis of soybean responses to drought stress along the day. *PLoS One* 10:e0139051
- Marin E, Nussaume L, Quesada A, Gonneau M, Sotta B, Huguency P, Frey A, Marion-Poll A (1996) Molecular identification of zeaxanthin epoxidase of *Nicotiana plumbaginifolia*, a gene involved in abscisic acid biosynthesis and corresponding to the ABA locus of *Arabidopsis thaliana*. *EMBO J* 15:2331–2342
- Maruyama K, Todaka D, Mizoi J, Yoshida T, Kidokoro S, Matsukura S, Takasaki H, Sakurai T, Yamamoto YY, Yoshiwara K, Kojima M, Sakakibara H, Shinozaki K, Yamaguchi-Shinozaki K (2012) Identification of cis-acting promoter elements in cold- and dehydration-induced transcriptional pathways in *Arabidopsis*, rice, and soybean. *DNA Res* 19:37–49
- Mizoi J, Ohori T, Moriwaki T, Kidokoro S, Todaka D, Maruyama K, Kusakabe K, Osakabe Y, Shinozaki K, Yamaguchi-Shinozaki K (2013) GmDREB2A;2, a canonical DEHYDRATION-RESPONSIVE ELEMENT-BINDING PROTEIN2-type transcription factor in soybean, is posttranslationally regulated and mediates

- dehydration-responsive element-dependent gene expression. *Plant Physiol* 161:346–361
- Nakashima K, Yamaguchi-Shinozaki K, 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
- Neves-Borges AC, Guimarães-Dias F, Cruz F, Mesquita RO, Nepomuceno AL, Romano E, Loureiro ME, de Fátima Grossi-de-Sá M, Alves-Ferreira M (2012) Expression pattern of drought stress marker genes in soybean roots under two water deficit systems. *Genet Mol Biol* 35:212–221
- Olsson A, Engstrom P, Soderman E (2004) The homeobox genes ATHB12 and ATHB7 encode potential regulators of growth in response to water deficit in Arabidopsis. *Plant Mol Biol* 55:663–677
- Osoerio MB, Bucker-Neto L, Castilhos G, Turchetto-Zolet AC, Wiebke-Strohm B, Bodanese-Zanettini MH, Margis-Pinheiro M (2012) Identification and in silico characterization of soybean trihelix-GT and bHLH transcription factors involved in stress responses. *Genet Mol Biol* 35:233–246
- Pedrosa AM, Cidade LC, Martins CP, Macedo AF, Neves DM, Gomes FP, Floh EI, Costa MG (2017) Effect of overexpression of citrus 9-cis-epoxycarotenoid dioxygenase 3 (CsNCED3) on the physiological response to drought stress in transgenic tobacco. *Genet Mol Res* 16:1–10
- Saito S, Hirai N, Matsumoto C, Ohgashi H, Ohta D, Sakata K, Mizutani M (2004) Arabidopsis CYP707As encode (+)-abscisic acid 8'-hydroxylase, a key enzyme in the oxidative catabolism of abscisic acid. *Plant Physiol* 134:1439–1449
- Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C_T method. *Nat Protoc* 3:1101–1108
- Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, Hyten DL, Song Q, Thelen JJ, Cheng J, Xu D, Hellsten U, May GD, Yu Y, Sakurai T, Umezawa T, Bhattacharyya MK, Sandhu D, Valliyodan B, Lindquist E, Peto M, Grant D, Shu S, Goodstein D, Barry K, Futrell-Griggs M, Abernathy B, Du J, Tian Z, Zhu L, Gill N, Joshi T, Libault M, Sethuraman A, Zhang XC, Shinozaki K, Nguyen HT, Wing RA, Cregan P, Specht J, Grimwood J, Rokhsar D, Stacey G, Shoemaker RC, Jackson SA (2010) Genome sequence of the palaeopolyploid soybean. *Nature* 463:178–183
- Schwartz SH, Tan BC, Gage DA, Zeevaart JAD, McCarty DR (1997) Specific oxidative cleavage of carotenoid by VP14 of maize. *Science* 276:1872–1874
- Seki M, Umezawa T, Urano K, Shinozaki K (2007) Regulatory metabolic networks in drought stress responses. *Curr Opin Plant Biol* 10:296–302
- Shin JH, Vaughn JN, Abdel-Haleem H, Chavarro C, Abernathy B, Kim KD, Jackson SA, Li Z (2015) Transcriptomic changes due to water deficit define a general soybean response and accession-specific pathways for drought avoidance. *BMC Plant Biol* 15:26
- Shinozaki K, Yamaguchi-Shinozaki K (2007) Gene networks involved in drought stress response and tolerance. *J Exp Bot* 58:221–227
- Shivhare R, Lata C (2017) Exploration of genetic and genomic resources for abiotic and biotic stress tolerance in pearl millet. *Front Plant Sci* 7:2069
- Singh AK, Sharma V, Pal AK, Acharya V, Ahuja PS (2013) Genome-wide organization and expression profiling of the NAC transcription factor family in potato (*Solanum tuberosum* L.). *DNA Res* 20:403–423
- Son O, Hur YS, Kim YK, Lee HJ, Kim S, Kim MR, Nam KH, Lee MS, Kim BY, Park J, Park J, Lee SC, Hanada A, Yamaguchi S, Lee IJ, Kim SK, Yun DJ, Söderman E, Cheon CI (2010) ATHB12, an ABA-inducible homeodomain-leucine zipper (HD-zip) protein of Arabidopsis, negatively regulates the growth of the inflorescence stem by decreasing the expression of a gibberellin 20-oxidase gene. *Plant Cell Physiol* 51:1537–1547
- Song L, Prince S, Valliyodan B, Joshi T, Maldonado dos Santos JV, Wang J, Lin L, Wan J, Wang Y, Xu D, Nguyen HT (2016) Genome-wide transcriptome analysis of soybean primary root under varying water-deficit conditions. *BMC Genomics* 17:57
- Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J* 37:914–939
- Tran LS, Nguyen HT (2009) Future biotechnology of legume. In: Emerich DW, Krishnan HB (eds) Nitrogen fixation in crop production. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, Madison, WI, pp 265–308
- Tripathi P, Rabara RC, Reese RN, Miller MA, Rohila JS, Subramanian S, Shen QJ, Morandi D, Bücking H, Shulaev V, Rushton PJ (2016) A toolbox of genes, proteins, metabolites and promoters for improving drought tolerance in soybean includes the metabolite coumestrol and stomatal development genes. *BMC Genomics* 17:102
- Valdés AE, Overnäs E, Johansson H, Rada-Iglesias A, Engström P (2012) The homeodomain-leucine zipper (HD-zip) class I transcription factors ATHB7 and ATHB12 modulate abscisic acid signalling by regulating protein phosphatase 2C and abscisic acid receptor gene activities. *Plant Mol Biol* 80:405–418
- Valliyodan B, Ye H, Song L, Murphy M, Shannon JG, Nguyen HT (2016) Genetic diversity and genomic strategies for improving drought and waterlogging tolerance in soybeans. *J Exp Bot* pii:erw433
- Wang YX, Liu ZW, Wu ZJ, Li H, Zhuang J (2016) Transcriptome-wide identification and expression analysis of the NAC gene family in tea plant [*Camellia sinensis* (L.) O. Kuntze]. *PLoS One* 11:e0166727
- Weeks DP, Spalding MH, Yang B (2016) Use of designer nucleases for targeted gene and genome editing in plants. *Plant Biotechnol J* 14:483–495
- Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY, Wei L (2011) KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* 39(Web Server issue):W316–W322
- Xu P, Cai W (2017) Functional characterization of the BnNCED3 gene in *Brassica napus*. *Plant Sci* 256:16–24
- Xu B, Ohtani M, Yamaguchi M, Toyooka K, Wakazaki M, Sato M, Kubo M, Nakano Y, Sano R, Hiwatashi YM, Kurata T, Yoneda A, Kato K, Hasebe M, Demura T (2014) Contribution of NAC transcription factors to plant adaptation to land. *Science* 343:1505–1508
- Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L, Wang J (2006) WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res* 34(Web Server issue):W293–W297
- Ye H, Liu S, Tang B, Chen J, Xie Z, Nolan TM, Jiang H, Guo H, Lin HY, Li L, Wang Y, Tong H, Zhang M, Chu C, Li Z, Aluru M, Aluru S, Schnable PS, Yin Y (2017) RD26 mediates crosstalk between drought and brassinosteroid signalling pathways. *Nat Commun* 8:14573
- Ying L, Chen H, Cai W (2014) BnNAC485 is involved in abiotic stress responses and flowering time in *Brassica napus*. *Plant Physiol Biochem* 79:77–87
- Zhang XM, Yu HJ, Sun C, Deng J, Zhang X, Liu P, Li YY, Li Q, Jiang WJ (2017) Genome-wide characterization and expression profiling of the NAC genes under abiotic stresses in *Cucumis sativus*. *Plant Physiol Biochem* 113:89–109
- Zheng Y, Huang Y, Xian W, Wang J, Liao H (2012) Identification and expression analysis of the Glycine max CYP707A gene family in response to drought and salt stresses. *Ann Bot* 103:743–756
- Zhong R, Demura T, Ye ZH (2006) SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. *Plant Cell* 18:3158–3170
- Zhong R, Richardson EA, Ye ZH (2007) Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis. *Planta* 225:1603–1611
- Zhu JK (2002) Salt and drought stress signal transduction in plants. *Annu Rev Plant Biol* 53:247–273
- Zhu JK (2016) Abiotic stress signaling and responses in plants. *Cell* 167:313–324