Genome-Wide Expression Profiling of Soybean Two-Component System Genes in Soybean Root and Shoot Tissues under Dehydration Stress

DUNG TIEN Le, RIE Nishiyama, YASUKO Watanabe, KEICHI Mochida, KAZUKO Yamaguchi-Shinozaki, KAZUO Shinozaki, and LAM-SON PHAN Tran

Signaling Pathway Research Unit, Plant Science Center, RIKEN Yokohama Institute, 1-7-22, Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan; Gene Discovery Research Group, Plant Science Center, RIKEN Yokohama Institute, 1-7-22, Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan; Agricultural Genetics Institute, Vietnamese Academy of Agricultural Science, Pham-Van-Dong Str., Hanoi, Vietnam and Japan International Center of Agricultural Sciences, Ibaraki 305-8686, Japan

*To whom correspondence should be addressed. Tel. +81 45-503-9593. Fax. +81 45-503-9591. E-mail: tran@psc.riken.jp

Edited by Satoshi Tabata
(Received 24 August 2010; accepted 29 November 2010)

Abstract

Two-component systems (TCSs) play vital functions in the adaptation of plants to environmental stresses. To identify soybean TCS genes involved in the regulation of drought stress response, we performed tissue-specific expression profiling of all 83 putative TCS genes in plants subjected to dehydration. Under well-watered conditions, the majority of soybean TCS genes were expressed higher in the root tissues. Additionally, a high variability in transcript abundance was observed for the TCS genes in both roots and shoots. Under dehydration, TCS genes were more responsive in shoots than in roots. Further analysis indicated that 50% more TCS genes were repressed by dehydration than induced. Specifically, 18 genes were induced by 2-fold or more, whereas 33 genes were down-regulated at least 2-fold by dehydration. TCS genes putatively involved in cytokinin and ethylene signalings strongly responded to dehydration, suggesting that crosstalk exists between different hormonal and stress pathways. Our study provides the first glance into the complex regulatory roles of soybean TCSs underlying their functions in response to dehydration. Additionally, these systematic expression analyses identified excellent dehydration-responsive candidate genes to further clarify soybean TCS functions in drought response and to enable the development of improved drought tolerance in transgenic soybeans.

Key words: soybean; two-component system; dehydration; expression profiling; RT-qPCR

1. Introduction

Drought stress is one of the several adverse environmental factors that are commonly encountered by plants and can result in significant reductions in crop productivity worldwide. In response to drought stress, plants activate a number of endogenous defence mechanisms that function to increase the drought tolerance. Plants have evolved molecular machinery that initiates complex signal-transduction networks which minimize the impact of a suboptimal water supply on plants. The identification of signalling pathways acting in stress-affected cells and the mutual interactions between these pathways are the major research efforts. As a result of the complexities of stress-related signal transduction and its huge potential for impact on modern agriculture, this research field has garnered a substantial amount of attention. The early events of plant adaptation to environmental stresses involve the perception of stress signals and subsequent signal transduction, leading to the activation of various physiological and metabolic responses.
In the signal-transduction networks involved in the perception of stress signals to stress-responsive gene expression, phosphorylation, which is mediated by two-component systems (TCSs) or His-to-Asp phosphorelays, is a key mechanism for stress signal transduction in cells. TCSs, consisting of sensor hybrid histidine kinases (HKs), histidine phosphotransfers (HPts) and effector response regulators (RRs), have been systematically identified and analysed in two completely sequenced and well-annotated model plant species: Arabidopsis thaliana and rice (Oryza sativa). A number of studies in Arabidopsis have suggested that among the Arabidopsis HKs (AHKs), AHK1, AHK2, AHK3 and AHK4 function in response to drought stress. In planta studies have demonstrated that the cytokinin (CK)-independent AHK1 functions as a positive regulator, whereas the CK-responsive AHK2, AHK3 and AHK4 function as negative regulators in drought stress signalling in both ABA-dependent and ABA-independent pathways. Although the AHK2, AHK3 and AHK4 proteins all positively regulate shoot growth, they exert a negative regulation of root growth. AHK1 also positively regulates shoot growth but its function in root growth is not known at this time. The involvement of AHKs in drought stress response suggests that the downstream Arabidopsis HPts (AHPs) and Arabidopsis RRs (ARRs) may also function in drought stress response. However, at the present time, there is not any in planta evidence or expression data, which have characterized the regulatory roles of AHPs in drought stress signalling. As for the ARRs, which can be classified into the type-A, type-B, type-C and pseudo ARRs based on their sequence signatures, loss-of-function studies indicated that among the type-A ARRs, ARR3, ARR4, ARR5 and ARR6 may function as positive regulators, whereas ARR8 and ARR9 as negative regulators in osmotic stress response, suggesting that these type-A ARRs might play a role in drought stress response. Recently, mutations in the pseudo aprr5, aprr7 and aprr9 genes enhanced tolerance of the triple mutant to drought stress, demonstrating that these three APRRs play negative roles in drought response. Since the rice TCS genes have been identified, expression studies have indicated that transcription of several rice TCS members, including HK, HPt and RR encoding genes, is altered by salt stress treatments. However, drought-related functions and/or expression profiling of rice TCS genes under dehydration and/or drought stress remain to be determined. TCSs were also identified in the important model legume Lotus japonicus. None of the abiotic stress-related data are currently available for L. japonicus TCS components.

Global soybean (Glycine max) production is frequently impacted by drought stress, which may reduce soybean yield by ~40%. Given the importance of TCS signalling pathways in the regulation of various biological processes and responses to environmental stimuli, including drought stress, we have recently compiled a list of putative TCS-associated components in soybean. Within the soybean genome, a total of 21 HK, 10 authentic and 3 pseudo HPt, 18 type-A RR, 15 type-B RR, 3 type-C RR and 13 pseudo RR encoding genes were identified. Comparative analysis of the Arabidopsis, rice and soybean TCS members revealed the conserved architecture of the TCSs between these plants. In this study, we aimed to identify tissue-related and dehydration-related soybean TCS genes by characterizing their expression profiles in both root and shoot tissues of soybean plants treated with a time-coursed dehydration stress. These findings will enable us to perform in planta functional analyses of the candidates and will allow us to identify appropriate stress-responsive TCS candidate genes and their respective promoters for the future improvement of drought resistance in soybean via genetic engineering. Therefore, the identification, characterization and molecular tailoring of novel TCS members will have the potential to overcome a number of important limitations involved in the generation of transgenic soybean plants with superior yield under drought conditions.

2. Materials and methods

2.1. Plant growth, dehydration treatment and collection of tissues

Soybean cv. Williams 82 seeds were germinated in 6-l pots containing vermiculite and were well-watered and grown under greenhouse conditions (continuous 30°C temperature, photoperiod of 12 h/12 h, 80 μmol m⁻² s⁻¹ photon flux density and 60% relative humidity). For tissue-specific expression profiling of TCS genes, root and shoot tissues were collected from 12-day-old soybean plants in three biological repeats. For expression profiling of TCS genes under dehydration stress, the dehydration treatment was carried out in time-course experiments to identify dynamic changes in transcripts in response to dehydration stress. Specifically, 12-day-old plants were carefully removed from soil, and roots were gently washed to remove soil. The plants were subsequently transferred onto a filter paper and allowed to dry for 2 and 10 h under the following conditions: 60% relative humidity, 25°C temperature and 10 μmol m⁻² s⁻¹ photon flux light intensity. The intensity of the dehydration stress treatment was quantified by the levels of relative water contents and standardized water contents, which were determined as described previously (n = 5;
Supplementary Fig. S1). For mock control, seedlings were removed from soil and washed exactly as above, then grown hydroponically in water for 2 and 10 h under the same conditions. Root and shoot tissues of dehydrated and control plants were separately collected in three biological repeats for expression profiling.

2.2. RNA isolation, DNase treatment and cDNA synthesis

Plant tissue samples were ground into a fine powder using a mortar and pestle. Total RNA was isolated using the TRIZOL reagent (Invitrogen), and RNA concentration was determined using the NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). For each sample, 4 μg of total RNA was digested according to the manufacturer's instructions in a 25-μl volume with Turbo DNA-free DNasel to remove genomic DNA contamination (Ambion, Austin, TX, USA). After DNasel treatment, RNA concentration was determined again with the NanoDrop spectrophotometer. The A_{260}/A_{280} values of all RNA samples used in this study were 2.08 or higher, and the A_{260}/A_{230} ratios of all samples were higher than 2.2. First-strand cDNA synthesis was performed using 1 μg of DNasel-treated RNA with the ReverTra Ace qPCR RT Kit (Toyobo, Japan) in a 20-μl reaction volume according to the manufacturer's supplied protocol.

2.3. Quantitative real-time PCR

Gene-specific primers for soybean TCS genes were designed using the Primer3 software. Primer specificity was confirmed by blasting each primer sequence against the soybean genome (Glyma1 model) followed by analysing the melting curves and amplicon fragments. Primers were redesigned if the corresponding melting curve did not yield a single sharp peak and/or if they had an electrophoresis pattern that failed to produce a single amplicon of the correct predicted length. The CYP2 gene was selected as a reference gene in the expression profiling of soybean genes as recommended previously. Quantitative real-time PCR (RT-qPCR) were performed in 96-well plates on a Stratagene MX3000P system (Agilent Technologies, Santa Clara, CA, USA) using Thunderbird™ SYBR® qPCR Mix (Toyobo, Japan) reagents. Primer sets of 0.4 μM final concentrations for each primer were used in a final volume of 10 μl well⁻¹. The thermal profile of the RT-qPCRs was at 95°C for 1 min, 40 cycles at 95°C for 15 s and at 60°C for 1 min. Dissociation curves were obtained using a thermal melting profile performed after the last PCR cycle: 95°C for 15 s followed by a constant increase in the temperature between 60°C and 95°C. Background-corrected raw fluorescence data were exported from the MX3000P system and analysed in LinRegPCR software with a built-in baseline correction and amplification efficiency calculation. Amplicon-based fluorescence thresholds were used to obtain the C_t values, and these values together with the amplicon-based mean efficiency were used for calculating the initial quantity of mRNA transcripts. Finally, the mRNA levels of each transcript were normalized with those of the corresponding CYP2 transcript.

2.4. Statistical analysis of the data

Mean values of three biological repeats were used to plot figures, and error bars on each figure represent the standard errors. When appropriate, a Student's t-test (one tail, unpaired, equal variance) was used to determine the statistical significance of the differential expression patterns between tissues and/or between treatments. Differential expression data which failed to pass the t-test with a P-value < 0.05 were regarded as insignificant. These insignificant data were regarded as ‘ubiquitous’ in the case of tissue-specific comparisons and as ‘unresponsive’ in comparisons between normal and stress treatments; regardless of the fold change between the expression levels. Dehydration-responsive genes were defined as the differentially expressed genes with at least 2-fold induction or repression at 2 h and/or 10 h after dehydration treatment, whereas their expression in the mock control seedlings did not significantly change in a similar manner (Student’s t-test, P < 0.05).

3. Results and discussion

3.1. Confirmation of primer specificity

Since the specificity of primers is crucial in RT-qPCR assays, we carefully designed each primer pair as described in the ‘Materials and methods’ section. Furthermore, we determined their specificity by analysing their respective melting curves and amplicon fragments. All primer sets used in this study produced only one peak in their respective melting curves, and corresponding amplicons detected under our RT-qPCR conditions generated a single band of expected size on 2% agarose gels (data not shown). The sequences and calculated amplification efficiencies of the specific primers used in this study can be obtained in Supplementary Table S1.

3.2. Expression of the soybean TCS genes in root and shoot tissues

Increasing evidence suggests that the mechanisms of drought resistance are either related to root and/or shoot traits. Root morphology and development is one of the key traits that is correlated with mechanisms of drought resistance. Because the distribution
of water within the rhizosphere is critical to maintaining function in different environmental conditions, plasticity for root traits is a vital factor to acquire resources. For instance, plants can adapt to drought by developing a longer taproot which helps reach lower soil layers where water is more readily available. In addition, an extensive fibrous root system can be useful for foraging subsoil surface moisture and nutrients such as phosphorus. On the other hand, a restraint of shoot growth has been shown to be advantageous in adverse environments by minimizing the evaporative leaf surface area. The growth restraint conferred by DELL proteins, whose functions are regulated by different hormones such as gibberellin (GA), auxin and ethylene, is beneficial and promotes survival. Overexpression of the GhDREB1 gene from cotton (Gossypium hirsutum) in Arabidopsis resulted in a cold, salt and osmotic stress-tolerant phenotype as a consequence of growth retardation caused by GA deficiency, decreased sensitivity to CKs and repression of CK signalling. Therefore, the appropriate control of shoot- and root-related morphological traits is a promising approach for developing drought resistance in a number of crops, including soybean.

In Arabidopsis, numerous reports have suggested that TCS members regulate shoot and root growth and branching. In order to identify candidate genes that could be potentially used for enhancing drought resistance by altering shoot and/or root growth when overexpressed or repressed in transgenic plant systems, we determined expression profiles for all root- and shoot-specific soybean TCS members. As shown in Fig. 1 and summarized in Table 1 (detailed in Supplementary Table S2), the majority of soybean TCS genes were preferentially or specifically expressed in roots according to the criteria defined for the analysis of tissue-specific expression (Table 1; Supplementary Table S2). Forty-eight of the 83 soybean TCS genes were expressed highly in roots. Among these 48 genes, 20 genes were further classified into a root-specific group, as their transcript abundance in roots was more than 10-fold higher than that of shoots. The remaining 28 genes fell into the root-preferential group as their corresponding root/shoot expression ratios were between 3 and 10 (Fig. 1, Table 1). One gene (GmRR18) was found to be ‘very specifically expressed’ in root tissues as its root/shoot transcript abundance ratio was more than 118-fold (Fig. 1C). Another gene (GmHK02) was identified as a root ‘exclusively expressed’ gene since it was only detected in root tissues (Fig. 1A). Only four genes (GmHP07, GmRR04, GmRR17 and GmPRR46) displayed higher expression in shoots and the remaining 30 genes showed ubiquitous expression patterns. All four of the shoot-related genes were grouped into the shoot-preferential group with root/shoot expression ratios ranging between −10 and −3 (Fig. 1, Table 1, Supplementary Table S2).

The expression levels of the soybean TCS genes were widely divergent. For example, the GmHK genes, which express more highly in roots than in shoots, can be divided into three groups (low, medium and high) based on their transcript abundance, in which the lowest (GmHK08) and highest (GmERS2) transcript abundance in roots was more than 3200-fold different (Fig. 1A). Among the GmHK genes, the genes encoding ethylene-receptor HKs, especially the GmERS1 and GmERS2, were expressed more highly than the others in both root and shoot tissues. Three GmHK genes (GmHK07–09), encoding proteins with highest homology to the osmosensor AHK1 of Arabidopsis, have a similar root-enriched expression pattern to the AHK1 gene. Similarly, TCS genes, coding for the Arabidopsis AHK4-like HKs in soybean (GmHK14–17), are more preferentially expressed in roots than in shoots (Fig. 1A), suggesting that a correlation between sequence conservation and expression patterns may exist. The soybean Hpt encoding TCS genes formed a highly divergent group based on their transcript abundance. Specifically, their highest/lowest expression ratios in root and shoot tissues were more than 16,000- and 65,000-fold, respectively (Fig. 1B). Among the GmRR genes, the type-A GmRRs possessed the most diverse expression levels in both roots and shoots (Fig. 1C). GmRR06 had the highest expression in the root tissues, which was 170,000-fold higher than that of the lowest one (GmRR16). The expression levels among type-A GmRR genes in shoots were less variable than in roots, and the difference in expression was 28,000-fold between the highest (GmRR10) and the lowest (GmRR16) expressed genes (Fig. 1C). Twelve of 15 genes encoding type-B GmRRs were expressed either preferentially (eight genes) or specifically (four genes) in root tissues (Table 1). There were no shoot-preferential expression patterns detected among the type-B GmRR genes (Fig. 1D). The greatest amount of divergence in root/shoot expression levels among GmRRs was found for GmRR27, which had a root/shoot transcript abundance ratio of 39-fold. According to a comparative sequence analysis, the 15 type-B GmRRs showed the highest amino acid sequence identity to five ARRs (ARR1, ARR2, ARR11, ARR12 and ARR14). With the exception of the AR14 gene that was not expressed in roots, semi-quantitative analyses determined that the remaining four ARR genes were expressed ubiquitously in root and leaf tissues. All of the GmPRR genes were expressed in both root and shoot tissues, of which four, eight and one genes were expressed root
Figure 1. Expression of soybean TCS genes in root (white bars) and shoot (black bars) tissues under normal conditions. (A) GmHK genes. (B) GmHP and GmPHP genes. (C) Type-A GmRR genes. (D) Type-B GmRR genes. (E) Type-C GmRR genes. (F) Pseudo GmPRR genes. Data represent the means and standard errors of three independent biological samples. Asterisks on the top of bars indicate statistically significant differences between tissues with a P-value < 0.05 (*) or 0.01 (**).
Table 1. Expression of the soybean TCS genes in root and shoot tissues

<table>
<thead>
<tr>
<th>Fold difference&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Root-specific (X &gt; 10-fold)</th>
<th>Root-preferential (3 &lt; X &lt; 10)</th>
<th>Ubiquitous&lt;sup&gt;b&lt;/sup&gt; (-3 &lt; X &lt; 3)</th>
<th>Shoot-preferential (-10 &lt; X &lt; -3)</th>
<th>Shoot-specific (X &lt; -10)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HKs</td>
<td>8</td>
<td>7</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Phosphotransfer proteins</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>RRs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type-A</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Type-B</td>
<td>4</td>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Type-C</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pseudo</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>28</td>
<td>30</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

<sup>a</sup>Fold difference was calculated as the ratio of mean expression levels of the same gene in roots and the shoots, when the root/shoot ratio is < 1.0, the ratio was reversed and a minus sign (−) was added.

<sup>b</sup>Genes whose differential expressions between root and the shoot tissues did not pass Student’s t-test with P-value < 0.05 are also classified as ubiquitously expressed genes.

preferentially, ubiquitously and shoot preferentially, respectively (Table 1, Fig. 1F). Among the soybean TCS genes, the three type-C GmRR genes (GmRR34–36) exhibited the lowest levels of expression under our experimental conditions (Fig. 1E). In an analysis of soybean short transcript sequence reads by Libault et al.,<sup>33</sup> the transcripts of two type-C GmRR genes (GmRR34 and GmRR35) were not detected in all the tissues except from the green pods. It should also be noted that the transcripts of GmHK05, GmHK06 and GmRR36 were not found in any of the nine tissues tested in their studies, including a transcriptomic analysis of root hair cells.<sup>33,36</sup> However, in our experimental settings, the presence of the transcripts of these genes were reliably detected, albeit at low levels (Fig. 1A and E). It is possible that this discrepancy is due to the duplicated nature of the soybean genome,<sup>37</sup> resulting in very high homology between duplicates. As a result, the method used in the authors’ study failed to detect expression because it could not assign reads matching with two or more closely related loci.<sup>33,36</sup>

3.3. Expression of the soybean TCS gene during dehydration

Plant TCSs have been implicated in the regulation of environmental stress responses, including drought stress,<sup>8,16–18,38</sup> suggesting that the TCS genes, which function in stress responses, may be used to enhance drought tolerance in plants via genetic engineering.<sup>8,38</sup> Stress-related soybean TCS genes can be predicted based on comparative sequence analyses<sup>21</sup>; however, this approach has limitations. Although TCSs have been systematically identified and characterized in Arabidopsis, rice, L. japonicus and soybean,<sup>5–7,19</sup> only a few members of Arabidopsis TCS genes have been identified as drought-related genes based on expression and/or functional analyses.<sup>8,38</sup> Cis-element-based targeted gene finding approach can also be used to predict stress-responsive genes, and tissue-specific genes.<sup>21,39–41</sup> However, because of their short length (5- to 9-bp core) and flexibility, the frequency of a cis-motif sequence in the whole genome is relatively high. In addition, a number of cis-element sequences might be syntactically correct without providing practical regulatory function.<sup>42</sup>

Therefore, our major goal of this study was to characterize the expression profile of all soybean TCS members under dehydration stress in order to precisely identify dehydration-responsive TCS genes that can potentially used to engineer soybean plants with improved drought resistance. We employed RT-qPCR analyses to conduct systematic expression profiling of all 83 TCS genes in root and shoot tissues of 12-day-old soybean plants subjected to 2 and 10 h dehydration stress. The evaluation of expression patterns in individual stressed tissues, rather than whole plants, might provide information on the mode of action of stress-responsive genes in specific tissues.<sup>43</sup>

As presented in Figs 2 and 3, a large number of soybean TCS genes appeared to be dehydration responsive. To precisely determine the dehydration-responsive genes, the expression of those soybean TCS genes, whose mRNA levels were found to be changed under dehydration treatment, was also examined in the water-treated control plant samples (Supplementary Figs S2 and S3). As a result, among 83 soybean TCS genes, a total of 18 induced and 33 repressed genes were identified (Fig. 4, Table 2, Supplementary Table S3). Previously, a cis-motif-based prediction using 12 stress-responsive cis-motifs suggested that out of 83 soybean TCS genes, 30 genes might be dehydration responsive because these genes were found to
contain dehydration-responsive ABRE and/or MYBR and/or MYCR motifs in their promoter regions. It should be noticed that the ABRE and MYBR were discovered as dehydration-inducible cis-motifs, whereas MYCR was shown to act as both dehydration-inducible and dehydration-repressible regulatory motifs. We found that among the 30 predicted dehydration-responsive TCS genes, 11 genes were confirmed by our RT-qPCR analysis. Eight genes (GmRK10, GmRR01, GmRR02, GmRR25, GmRR34, GmRR35, GmRR36 and GmPRR39) are induced, and three genes (GmHP01, GmRR09 and GmRR29), which contain MYCR(s) in their promoter regions, are repressed (Supplementary Table S3). Thus, although the cis-

Figure 2. Expression of TCS genes encoding HKs and HPts in root (white bars) and shoot (black bars) tissues of soybean plants under dehydration stress. (A) Expression of TCS genes encoding HKs. (B) Expression of TCS genes encoding HPts. Relative gene expression levels were normalized to a value at 1 in the untreated plant samples (0 h). Data represent the means and standard errors of three independent biological samples. Asterisks on the top of bars indicate statistically significant differences when compared with 0 h with a P-value < 0.05 (*) or 0.01 (**).
Figure 3. Expression of TCS genes encoding RRs in root (white bars) and shoot (black bars) tissues of soybean plants under dehydration stress. (A) Type-A GmRR genes. (B) Type-B GmRR genes. (C) Type-C GmRR genes. (D) Pseudo GmPRR genes. Relative gene expression levels were normalized to a value at 1 in the untreated plant samples (0 h). Data represent the means and standard errors of three independent biological samples. Asterisks on the top of bars indicate statistically significant differences when compared with 0 h with a P-value < 0.05 (*) or 0.01 (**).
element-based targeted gene finding approach has demonstrated a wide application in the genome-wide prediction due to the availability of a large number of cis-elements in many plant species. Expression profiles of the cis-motif-based predicted genes should be verified experimentally using an expression profiling approach, such as RT-qPCR, prior to the launching of laborious in planta functional studies.\textsuperscript{39,40,42,45} Although the soybean TCS genes were predominantly expressed in roots than in shoots (Fig. 1), their expression patterns upon dehydration stress were opposite (Fig. 4). Specifically, there were only 8 and 17 genes whose expression was induced or repressed in the roots, respectively, upon dehydration. At the same time, there were 14 and 20 genes whose expression was either induced or repressed, respectively, in shoots (Fig. 4). Only a small number of the genes responded in a similar fashion (induced or repressed) in both tissues. Specifically, the GmHK06, GmHK08 and GmHK11 genes were down-regulated, and the GmRR16, GmRR34, GmRR35 and GmRR36 genes were up-regulated. These data support the hypothesis that the majority of soybean TCS genes respond to dehydration stress in either root- or shoot-specific manner. It is worth mentioning that the dehydration responsiveness of the GmHK10 gene were opposite in the root and the shoot tissues. The expression of GmHK10 was induced in the roots at 2 h but was repressed in the shoots at 10 h after dehydration treatment (Fig. 2A, Supplementary Fig. S2A). Not only having a larger number of dehydration-responsive genes, but the extent of the responsiveness of GmHKs, i.e. the fold change, was also higher in shoot tissues. For example, the expression levels of the GmHK01 and GmHK12 genes, which encode proteins most similar to the Arabidopsis CKI1 and AHK3, respectively,\textsuperscript{21} exhibited a 5–6-fold induction in shoots but insignificant induction in roots (Supplementary Table S3). Similarly, the extent of repression was also more severe in shoots than in roots. A maximum of 6-fold repression was found in roots, whereas a 17-fold repression was observed in shoots (Supplementary Table S3). GmHK08 was the most repressed among all soybean TCS genes during dehydration treatment and in both tissues, and its expression in shoots at 10 h after dehydration treatment was not detectable (Fig. 2A, Supplementary Fig. S2A). In Arabidopsis, the osmosensor AHK1 and the CK receptor AHK2, AHK3 and AHK4 have been shown to function in drought stress response. The expression of all four of these AHK genes is induced by dehydration.\textsuperscript{8} Among the three AHK1-like protein encoding GmHK07–09 genes,\textsuperscript{21} the expression of GmHK07 was induced by dehydration, suggesting that GmHK07 may play a positive regulatory role in drought response of soybean plants similar to that of AHK1 in Arabidopsis.\textsuperscript{8} Out of eight GmHK genes (GmHK10–17), which are predicted to encode CK-receptor HKs based on sequence similarity to their Arabidopsis counterparts,\textsuperscript{21} GmHK10 and GmHK12 showed significant induction in roots and shoots, respectively (Fig. 2A, Supplementary Fig. S2A). It is possible that GmHK10 and GmHK12 may act as negative

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
Family & Number of genes & Induced genes & Repressed genes \\
\hline
HKs & 21 & 5 & 10 \\
Phosphotransfers & & & \\
\hspace{1cm} Authentc & 10 & 2 & 2 \\
\hspace{1cm} Pseudo & 3 & 0 & 2 \\
RRs & & & \\
\hspace{1cm} Type-A & 18 & 4 & 7 \\
\hspace{1cm} Type-B & 15 & 2 & 8 \\
\hspace{1cm} Type-C & 3 & 3 & 0 \\
\hspace{1cm} Pseudo & 13 & 2 & 4 \\
Total genes & 83 & 18 & 33 \\
\hline
\end{tabular}
\caption{Genes with altered expression under dehydration in each family}
\end{table}
regulators in drought stress response in a similar fashion as their Arabidopsis orthologs. Based on our analyses, all of the GmHK07, GmHK10 and GmHK12 genes appear to be good candidates for in planta studies.

Among the 13 soybean TCS genes encoding HPt proteins, two genes (GmHP03 and GmHP06) were induced and four other genes were repressed upon dehydration (Table 2). Interestingly, none of the induced genes were found in root tissues. Among the induced genes, GmHP06 exhibited the highest induction (>6-fold) in shoots, while its induction in roots was not significant. The most repressed gene was GmPHP2, whose expression was reduced by 7.5-fold in shoots upon dehydration (Fig. 2B, Supplementary Fig. S2B, Supplementary Table S3).

The expression of soybean GmRR genes followed different patterns (Fig. 3A, Supplementary Fig. S3, Supplementary Table S3). Repression of the type-A GmRRs genes in roots was not statistically significant. In addition, only two genes (GmRR01 and GmRR16) were induced in the same tissues. In shoots, three and six genes of this group were induced and repressed upon dehydration, respectively. Together, 11 of the 18 type-A GmRR genes were either induced and/or repressed by dehydration in root and/or shoot tissues, suggesting their diverse functions in the regulation of dehydration stress response (Supplementary Table S3). Diverse functions were also reported for Arabidopsis type-A ARRs. Mutations in arr3, arr4, arr5 and arr6 increase sensitivity, but an additional loss of ARR8 and ARR9 decreases the sensitivity to osmotic stress. These data suggest that ARR3, ARR4, ARR5 and ARR6 may function as positive regulators, whereas ARR8 and ARR9 function as negative regulators. In soybean, the six dehydration-repressed type-A GmRR genes (GmRR07–09 and GmRR11–13) encode ARR8- and ARR9-like GmRRs, providing correlative evidence that these GmRRs function in stress response, and may act as negative regulators in a similar fashion as their ARR8 and ARR9 orthologues (Supplementary Table S3). Similarly, the ARR4- and ARR6-like protein encoding GmRR01 and GmRR02 genes were up-regulated in response to drought, suggesting that they may function as positive regulators in stress response in a similar fashion as their Arabidopsis counterparts. The expression of type-B GmRR genes in soybean, which encode transcription factors, was mostly repressed in both tissues. Only two genes (GmRR25 and GmRR32) were induced in response to drought, and they exhibited shoot specificity (Fig. 3B, Table 2, Supplementary Fig. S3, Supplementary Table S3). As for the type-C GmRR genes, all three genes (GmRR34–36) were found to be induced strongly in both roots and shoots by dehydration (Fig. 3C, Table 2, Supplementary Fig. S3, Supplementary Table S3).

Among 13 pseudo GmPRR genes, the expression of six genes was altered by dehydration, suggesting that they may play a physiological role in response to dehydration (Fig. 3D, Table 2, Supplementary Fig. S3, Supplementary Table S3). Genes encoding GmPRR39 and GmPRR44 were significantly induced in shoots and roots, respectively, upon dehydration treatment. Four other genes encoding GmPRRs were repressed upon dehydration, of which two genes (GmPRR46 and GmPRR47) were repressed in the roots, whereas two other genes (GmPRR48 and GmPRR49) were repressed in the shoots (Supplementary Table S3). The dehydration-inducible GmPRR39 and GmPRR44 genes were shown to encode either APRR5- or APRR9-like proteins. In Arabidopsis, the APRR5 and APRR9 proteins were reported to negatively regulate drought response. Therefore, it is possible that the dehydration-responsive GmPRR39 and GmPRR44 may function as negative regulators in soybean as well. On the other hand, although the transcripts of GmPRR38, GmPRR42, GmPRR43 and GmPRR45 were accumulated under dehydration, the transcripts of these genes were also increased in a similar manner in the water-treated mock control plant tissues (Supplementary Fig. S3), suggesting that circadian clock-related regulation affected the expression of these GmPRR genes rather than dehydration stress. This is not a surprise because it has been reported previously that these four GmPRR genes encode circadian clock-associated APRR1-, APRR7- and APRR9-orthologous proteins.

3.4. Clustering analysis of the expression data of soybean TCS genes

A gene’s expression is regulated through several mechanisms, of which some are still unknown. Nonetheless, the regulation of gene expression via the interaction of transcription factors and promoters has been well documented. Similarities in expression patterns and responsiveness to stresses may likely be a result of the involvement of similar cis-elements and/or transcription factors. In order to gain an overall understanding of the expression patterns of soybean TCS genes, we performed hierarchical clustering using their log-transformed expression data. Figure 5 shows the result of hierarchical analysis of soybean TCS genes performed on each of the groups (GmHKs, GmHPs and GmRRs). As shown in Fig. 5A, the expression patterns of GmHK genes clustered into two distinct nodes. Within these two distinct nodes, each group was separated into two subnodes. The high correlation coefficient (0.85–0.96) of each subnode indicates that the expression patterns of
the genes in each node were highly similar to one another. The expression of several genes with high homology, which is an indication of duplicated pairs, did not cluster in the analysis. For instance, GmHK02 and GmHK05, and GmHK10 and GmHK11 did not cluster, implying that their divergence in expression was the cause. The expression patterns of GmHP genes in roots and shoots during dehydration were clustered into three nodes with high correlation coefficient values (Fig. 5B). It should be noted that GmHP07 and GmHP08 did not fall into those nodes. With the exception of GmPHP1 and GmPHP2, the expressions of other highly homologous pairs in this group clustered together. These data indicated that functional redundancy is still high among duplicated Hpt proteins of soybean. Expression profiles of each of the RR types (A, B, C and pseudo) clustered into two nodes mainly based on the expression levels (Fig. 5C–E). Some of the groups were clustered into two or more subgroups based on the responsiveness to dehydration in the tested tissues (Fig. 5C–E).

To find out how the soybean TCS genes cluster in response to dehydration, we conducted a hierarchical analysis based on the log-transformed ratio of the expression values with and without stress treatment. The groups of genes clustered in this analysis will not take into account the gene-to-gene expression divergence. We found that several nodes were
formed (Supplementary Fig. S4), implying the possibility that similar dehydration-regulated pathways exist for each node.

3.5. Conclusions

This report has provided the first insight into the previously uncharacterized TCS members of soybean and placed a special emphasis on the relation to dehydration stress responsiveness. Our results have provided useful information by identifying candidate dehydration-responsive genes. By combining these genes with their associated dehydration-responsive promoters, scientists can utilize these resources to engineer soybean plants for enhanced stress resistance. We have selected a number of dehydration-responsive genes for further analysis in *Arabidopsis* to corroborate their functional significance *in planta*. By implementing these functional studies, we aim to identify a suite of candidate genes that are best suited for the genetic engineering of soybean plants with improved drought resistance.

Supplementary Data: Supplementary data are available at www.dnaresearch.oxfordjournals.org.

Funding

D.T.L. is supported by the RIKEN Foreign Postdoctoral Fellowship. This work was funded by a Start-up Support grant (M36-57000) from the RIKEN Yokohama Institute Director Discretionary Funds to L.-S.F.T.

References