The Auxin Response Factor Transcription Factor Family in Soybean: Genome-Wide Identification and Expression Analyses During Development and Water Stress

Chien Van Ha[1,2,3], Dung Tien Le[1,2], Rie Nishiyama[1], Yasuko Watanabe[1], Saad Sulieman[1], Uyen Thi Tran[1], Keiichi Mochida[4,5], Nguyen Van Dong[2], Kazuko Yamaguchi-Shinozaki[6], Kazuo Shinozaki[4,7], and Lam-Son Phan Tran[1,*]

Signaling Pathway Research Unit, RIKEN Center for Sustainable Resource Science, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan[1]; National Key Laboratory of Plant Cell Biotechnology, Agricultural Genetics Institute, Vietnamese Academy of Agricultural Science, Pham-Van-Dong Str., Hanoi, Vietnam[2]; Post-Graduate Program, Vietnamese Academy of Agricultural Science, Thanhtri, Hanoi, Vietnam[3]; Biomass Engineering Program Cooperation Division, RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan[4]; Kihara Institute for Biological Research, Yokohama City University, 641-12 Maioka-cho, Totsuka-ku, Yokohama, Kanagawa 244-0813, Japan[5]; Graduate School of Agriculture and Life Sciences, University of Tokyo, Tokyo 113-8657, Japan[6] and Gene Discovery Research Group, RIKEN Center for Sustainable Resource Science, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan[7]

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

Abstract

In plants, the auxin response factor (ARF) transcription factors play important roles in regulating diverse biological processes, including development, growth, cell division and responses to environmental stimuli. An exhaustive search of soybean genome revealed 51 *GmARF*s, many of which were formed by genome duplications. The typical GmARFs (43 members) contain a DNA-binding domain, an ARF domain and an auxin/indole acetic acid (AUX/IAA) dimerization domain, whereas the remaining eight members lack the dimerization domain. Phylogenetic analysis of the ARFs from soybean and *Arabidopsis* revealed both similarity and divergence between the two ARF families, as well as enabled us to predict the functions of the GmARFs. Using quantitative real-time polymerase chain reaction (qRT-PCR) and available soybean Affymetrix array and Illumina transcriptome sequence data, a comprehensive expression atlas of *GmARF* genes was obtained in various organs and tissues, providing useful information about their involvement in defining the precise nature of individual tissues. Furthermore, expression profiling using qRT-PCR and microarray data revealed many water stress-responsive *GmARF*s in soybean, albeit with different patterns depending on types of tissues and/or developmental stages. Our systematic analysis has identified excellent tissue-specific and/or stress-responsive candidate *GmARF* genes for in-depth *in planta* functional analyses, which would lead to potential applications in the development of genetically modified soybean cultivars with enhanced drought tolerance.

Key words: ARF transcription factor family; soybean; structural analysis; expression analysis; water stress

1. Introduction

Soybean [*Glycine max* (L.) Merrill] provides a major source of food and oil for human consumption, animal feed and bioenergy, and has capacity to fix atmospheric nitrogen through symbiosis.[1–4] Soybean growth, productivity and seed quality are adversely affected by a wide range of environmental stresses,

particularly drought which may reduce soybean yield by >40%.[5,6] To cope with drought stress, plants activate a number of defence mechanisms, including the perception of stress signals and subsequent signal transduction, leading to the activation of various physiological and metabolic responses.[7−10] Within the regulatory networks, various transcription factors (TFs) and *cis*-acting elements contained in stress-responsive promoters function as molecular switches for gene expression and terminal points of signal transduction in the regulatory processes. Increasing evidence suggests that TF-encoding genes have a great potential in genetic engineering of transgenic crops with stable yield under stress conditions.[11−15]

The phytohormone auxin has been known to regulate various aspects of plant growth and development.[16−21] Increasing evidence also suggests that auxin, either alone or together with other hormones, plays important roles in regulation of plant responses to environmental stimuli.[22−27] Expression profilings have revealed that many auxin-responsive genes are responsive to various abiotic stressors.[28−30] Later root development, which is one of the important drought-stress-related trait, was shown to be coordinately regulated by auxin, abscisic acid (ABA) and cytokinin through ABI4 (ABA INSENSITIVE 4) TF.[31] Numerous genetic and biochemical studies in *Arabidopsis* have provided evidence that transcriptional regulation of auxin response genes are regulated by two large TF families, the auxin response factor (ARF) and the Aux/IAA families.[32] In *Arabidopsis*, there are 23 ARFs most of which contain a conserved N-terminal DNA-binding domain (DBD), a variable middle transcriptional regulatory region (MR) and a carboxy-terminal dimerization domain (CTD).[33,34] The DBD of ARFs specifically binds to the conserved auxin response element (AuxRE, TGTCTC) in promoter regions of primary or early auxin-responsive genes. The structure of the transcriptional regulatory region (TRR) of each ARF determines whether the ARF acts as an activator or repressor. Activation domain (AD) of ARFs is usually enriched in glutamine (Q), serine (S) and leucine (L), while repression domain (RD) is enriched in either S, L and proline (P); S, L and/or glycine (G) or S. The ARF CTD is modular with amino acid sequence related to domains III and IV in Aux/IAA proteins, making it function as a dimerization domain among the ARF CTDs or with several Aux/IAA proteins.[32,35−37]

Given the importance of ARF TFs in diverse biological and physiological processes, and their potential applications for the development of improved stress-tolerant transgenic crop plants, the ARF TF families have been identified and characterized in a number of crop species, such as maize (*Zea mays*),[38,39] rice (*Oryza sativa*),[28,29,40] sorghum (*Sorghum bicolor*),[30] tomato (*Solanum lycopersicum*)[41] and Chinese cabbage (*Brassica rapa*).[42] The recent completion of genomic sequence of the model soybean cultivar Williams 82 (W82)[43] has enabled the soybean community to perform gene discovery in soybean with the aim to identify potential candidate genes for the improvement of yield under adverse environmental stress via genetic engineering.[44−49] In the present study, we carried out a genome-wide analysis of the soybean ARF family to identify all the putative GmARF TFs that were subsequently subjected to a phylogenetic analysis with their *Arabidopsis* counterparts to identify gene orthologs and clusters of orthologous groups, enabling functional prediction. We also performed a comprehensive expression analysis of all *GmARF* genes in various tissues using quantitative real-time polymerase chain reaction (qRT-PCR) or the wealth of available expression data, which were generated either by high-throughput microarray analyses or by Illumina transcriptome sequencing. These data, in turn, provided important complementary information to assist in the elucidation of the functions of the GmARFs. Since we have strong interest in research on mechanisms of water stress responses in soybean, we used a time-course dehydration stress treatment and subsequent qRT-PCR analysis as a precise mechanism to analyse the root- and shoot-related expression of all identified *GmARF* genes under normal and dehydration stress conditions. The results of this systematic qRT-PCR analysis have ultimately enabled us to identify appropriate root- or shoot-related and/or dehydration-responsive *GmARF* candidate genes for further *in planta* functional analyses towards biotechnological applications for the improvement of drought tolerance in soybean.

2. Materials and methods

2.1. Plant growth, treatments and collection of tissues

W82 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), as previously described.[15,50] Subsequently, root and shoot tissues were separately collected from 12-day-old soybean plants [vegetative cotyledon (VC) stage with unrolled unifoliolate leaves] in three biological replicates for tissue-specific expression profiling of *GmARF* genes. For expression profiling of *GmARF* genes under dehydration stress, the dehydration treatment was carried out in a time-course experiment as essentially described by Le *et al*.[48] Briefly, 12-day-old plants grown under well-watered conditions were carefully removed from pots and roots were gently washed to remove the soil. Subsequently, the plants were transferred onto a filter paper and allowed to dry for 0, 2 and 10 h. Root and

shoot tissues were collected separately in three biological replicates and were immediately frozen in liquid nitrogen until use.

2.2. Identification of the GmARF members in soybean

All GmARF TFs predicted in soybean were collected for manual analysis from various plant TF databases,[34,51−53] and only those GmARFs containing full open reading frames (ORFs), as predicted by Glyma v1.1 (http://www.phytozome.net/soybean), were used for further analyses. Genes with threshold of ≥90% nucleotide sequence identity were considered as duplicated genes.[54] Tandem duplicates were defined as duplicated genes located within 20 loci from each other.[55]

2.3. Phylogenetic analysis

Sequence alignments of all identified ARFs from *Arabidopsis* and soybean were performed with a gap open penalty of 10 and a gap extension penalty of 0.2 using ClustalW implemented on the MEGA 5 software.[56,57] The alignments were subsequently visualized using GeneDoc (http://www.nrbsc.org/gfx/genedoc/) as presented in Supplementary Fig. S1. The sequence alignments were also used to construct the unrooted phylogenetic tree by the neighbor-joining method using MEGA 5. The confidence level of monophyletic groups was estimated using a bootstrap analysis of 10 000 replicates. Only bootstrap values >50% are displayed next to the branch nodes.

2.4. Expression analyses of GmARF genes using microarray data and soybean Illumina expression data

For tissue-specific expression analysis of *GmARF* genes, microarray-based expression data for 68 types of tissues and organs housed in Genevestigator (https://www.genevestigator.com/) were used.[58] Illumina transcriptome sequencing data provided by Libault *et al.*[59,60] were also used to evaluate the expression of *GmARF* genes in eight tissues: nodules of 35-day-old soybean plants (harvested after 32 days of inoculation of the 3-day-old plants), 14-day-old shoot apical meristem (SAM), flowers (reproductive R2 stage), green pods (R6 stage), 18-day-old trifoliate leaves, roots (V2 stage), root tips and root hairs of 3-day-old seedlings.

For expression analysis of *GmARF* genes in soybean leaves at V6 and R2 stages under drought stress, which was imposed on the plants by withholding water from the pots until the volumetric soil moisture content reduced to <5%, microarray data recently published by Le *et al.*[61] were used. At the V6 stage, soybean plants had six unrolled trifoliate leaves and seven nodes, while at R2 full bloom stage, open flowers were found on any of the top two nodes on the main stem.

2.5. RNA isolation, DNaseI treatment and cDNA synthesis

Plant tissue samples were ground in liquid nitrogen using a mortar and pestle. Total RNA was isolated using the TRIZOL reagent according to the manufacturer's supplied protocol (Invitrogen). RNA concentration and integrity were measured prior to DNase I digestion with the NanoDrop UV-Vis spectrophotometer (NanoDrop Technologies). DNase I treatment and cDNA synthesis using Turbo DNA-free DNase I (Ambion) and the ReverTra Ace® qPCR RT Kit (Toyobo, Japan), respectively, were performed as previously described.[62]

2.6. qRT-PCR and statistical analysis of the data

Primers for qRT-PCR were designed as previously described (Supplementary Table S1).[49] Primer specificity was first confirmed by blasting each primer sequence against the soybean genome (Glyma v1.1).[43] qRT-PCR reactions and data analyses were performed according to previously published methods.[49] The *60s* gene was used as a reference gene as recommended by Le *et al.*,[62] and the delta-CT method was used to calculate the initial amount of target genes. When appropriate, 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. Considering the biological significance of the differential expression in this study, we adopted a cut-off value of 3-fold for tissue-specific expression, and 2-fold (at least at one time point) when analysing stress induction or repression. The expression levels were designated as 'tissue-specific,' 'induced' or 'repressed' only if such differences met the above criteria and passed the Student's *t*-test.

3. Results and discussion

3.1. Identification of the GmARF members in soybean

Currently, three databases, namely SoybeanTFDB,[51] SoyDB[53] and PlantTFDB,[34] provide access to the TF repertoire of soybean, which was obtained by genome-wide analysis of the Glyma v1.0 model. Interestingly, each group provided different numbers of the putative GmARF TFs in their databases. SoyDB reported the highest number of putative GmARFs (101), while SoybeanTFDB and PlantTFDB predicted only 75 and 55 GmARFs, respectively. As an initial step, we collected the sequences for all of the putative GmARFs from the three databases for sequence comparison to make a list of all the GmARF proteins. Because the Glyma v.1.1 has been available to public since July 2012 and no update has been reported yet by any of the above-mentioned databases, we blasted each GmARF protein sequence against the Glyma v.1.1 proteome using

blastp to identify putative all GmARF TFs that contain full ORFs by the Glyma1 v1.1 annotation. Thus, we were able to identify 51 GmARFs with annotated full ORF, and only these full-length (FL) GmARF TFs were used for further analyses. If Glyma v1.1 annotation predicted several splice variants for a given *GmARF* gene, all the alternative splice variants were carefully checked using soybean FL-cDNA information publicly available at http://rsoy.psc.riken.jp/. When FL-cDNA information is not available, splice variants that encode the longest ORFs were selected as representatives for subsequent sequence alignments. Supplementary Table S2 provided relevant information, including gene IDs as defined by the Glyma v1.1 model for each identified GmARF protein, lengths of amino acid sequences and corresponding available FL-cDNA accession numbers for all 51 GmARFs. A uniform nomenclature for all the *GmARF* genes identified in this work was adopted to facilitate scientific communication, taking into account the order of the chromosomes (Supplementary Table S2). Additionally, the cDNAs and protein sequences of all 51 GmARFs were also supplied in Supplementary Dataset 1 for convenient downloading and use.

3.2. Chromosomal distribution, structural and phylogenetic analyses of the GmARFs

To gain an insight into the genome organization of the *GmARF* genes, the position of each *GmARF* gene was obtained from Gbrowse (http://www.phytozome.net/cgi-bin/gbrowse/soybean/). The *GmARF* genes were found to be distributed on every chromosome in soybean (Fig. 1A), and the relative location of each of the *GmARF*s was illustrated on their respective chromosome (Fig. 1B). Chromosomes VI, IX, X, XIX and XX contain the lowest number of the *GmARF*s with only one member on each chromosome (∼ 2%), while chromosome XIII possesses the highest number of *GmARF*s with 7 of the 51 members (∼13%) (Fig. 1A and B).

Next, we were interested in identifying duplicated genes, because they represent the source of genetic materials for studying evolution and diversification.[63] Among 51 *GmARF* genes, we found 17 duplicates; each pair shares a ≥90% nucleotide sequence identity. On the basis of their physical localization, none of these duplicated pairs were found to be tandem duplicates as all pairs of the duplicated genes are located on different chromosomes (Fig. 1B). Evolutionary studies have suggested that the soybean genome experienced a tetra-ploidization event ∼10−15 million years ago and subsequently went through extensive gene rearrangements and deletions to become diploidized.[64,65] Since duplications resulting from whole-genome duplication events are largely retained,[43] we can observe in soybean that multigene families, such as TF-encoding and hormone

biosynthesis-related families,[50,65,66] contain highly related genes, making functional redundancy; a phenomenon that is common in plants.

The features and number of domains and subdomains present in the GmARF sequences provide useful information for the prediction of their functions.[32] Protein sequence alignment of the GmARFs with their *Arabidopsis* counterparts confirmed that all the GmARFs have a typical ARF-type structure with a conserved DBD that consists of a plant-specific B3-type subdomain and an ARF subdomain required for efficient *in vitro* binding to the AuxRE (Fig. 1C; Supplementary Fig. S1). Among the 51 GmARFs, which could be classified into six groups (Groups a–f) based on their structure, nine GmARFs (08, 21, 23, 26, 29, 30, 35, 38 and 51; Groups d, e and f) contain an additional short segment of 12−44 residues within their DBD. As for the CTD, eight members (GmARF08, 16, 30, 32, 34, 38, 41 and 51; Groups c, e and f) lack the CTD and the remainings have the typical CTD with both III and IV subdomains. Comparing with the ARF members identified in other dicot plants, soybean (15.68%) and *Arabidopsis* (17.39%) have similar percentage of CTD-truncated ARFs, while *B. rapa* and tomato have a higher rate of CTD-truncated ARFs with 22.58 and 28.57%, respectively.[32,41,42] With regard to the middle region (MR), of 51 GmARFs, 19 members contain the QSL-rich region (Group a), whereas the remaining GmARFs, except the GmARF51 (Group f), possess a TRR enriched in either SPL (Groups b and d), SLG (Group c) or S (Group e) (Fig. 1C; Supplementary Fig. S1). This difference in TRR signatures suggests that the GmARFs of the former group might act as an activator and those of the latter as repressors, respectively, based on the evidence accumulated from functional analyses of *Arabidopsis* ARFs.[32] In addition, similar to the typical AtARFs, all the GmARFs contain a conserved putative monopartite nuclear localization signals (NLS) at the end of the DBD (Supplementary Fig. S1).[66] This consensus monopartite NLS was also predicted in OsARFs of rice, which was recently shown to be able to direct the gene product into the nucleus by a synthetic green fluorescent protein fusion assay.[40]

As a means to classify subgroups and to identify the evolutionary relationships between GmARFs and their *Arabidopsis* ARF counterparts (AtARFs), a phylogenetic analysis of GmARFs and AtARFs was performed. The unrooted phylogenetic tree was built from the alignment of the FL-amino acid sequences of 51 GmARFs and 23 AtARFs. As shown in Fig. 2, all GmARFs and AtARFs were classified into four major groups based on their phylogenetic relationship. Group I could be further divided into two subgroups: Ia and Ib. Groups Ia, II and III contained GmARFs and AtARFs with relatively high sequence similarity, suggesting that the members of these subgroups derived from common
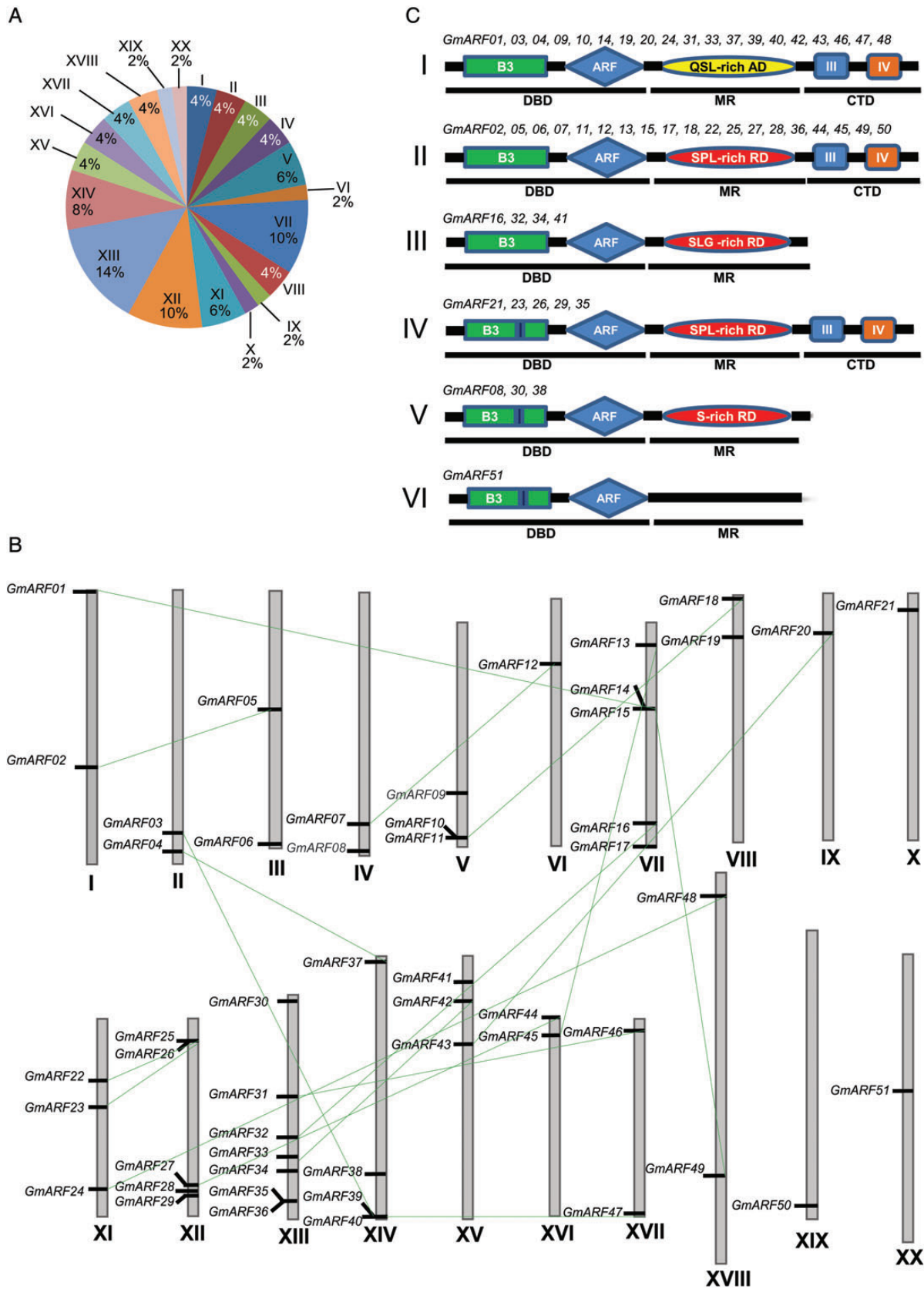
Figure 1. Chromosomal distribution of 51 soybean *GmARF* genes identified in this study and structural analysis of the GmARF proteins. (A) Chromosomal distribution of *GmARF* genes with indication of percentages of *GmARF*s located on each chromosome. (B) Graphical representation for chromosomal localization of *GmARF* genes. Greek numbers indicate chromosome numbers. (C) Graphical representation for domain organization of GmARF proteins. A typical ARF contains a DBD, which consists of a B3 subdomain and an auxin response (ARF) subdomain, a MR and a CTD.
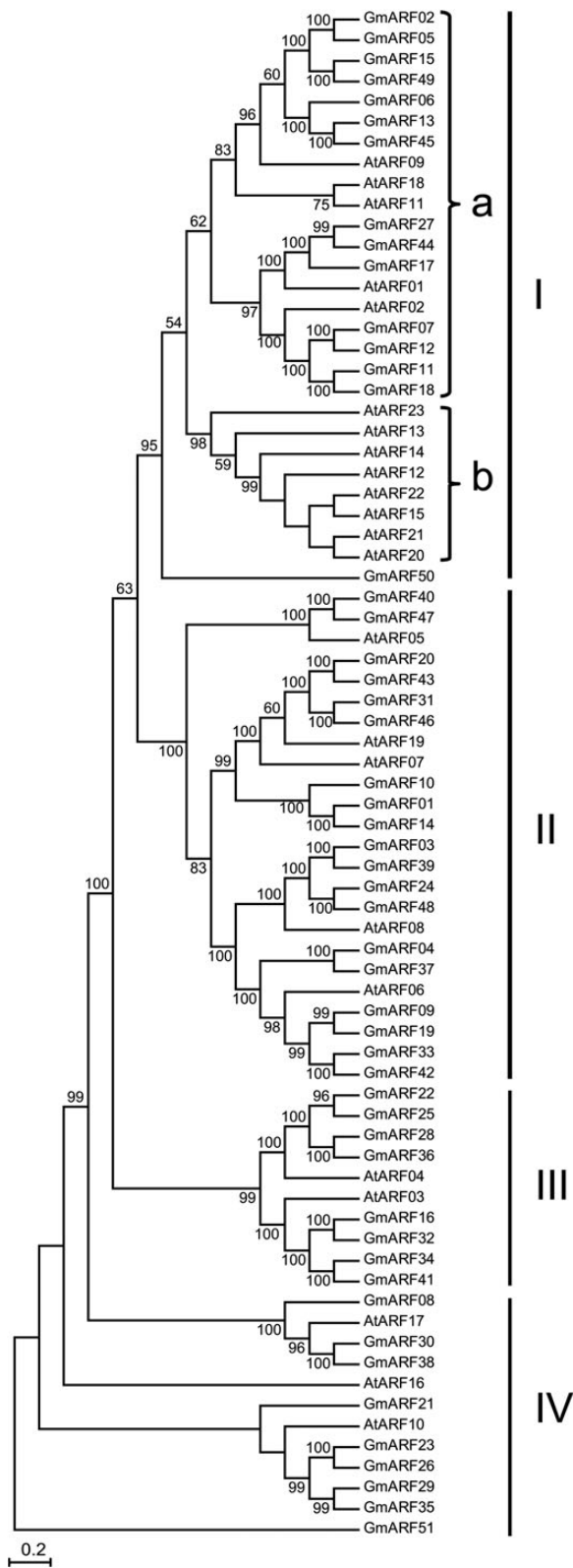
Figure 2. Phylogenetic relationship of ARFs from *Arabidopsis* and soybean. The unrooted phylogenetic tree was constructed using the full ORFs of ARF proteins. The bar indicates the relative divergence of the sequences examined. Bootstrap values >50% are displayed next to the branch.

ancestors. In addition, Group II was consisted of the GmARFs that possess QSL-rich region, thereby they might act as activators. On the other hand, Groups Ia and III contained GmARFs with SPL- or SLG-rich region, suggesting that these GmARFs might have repression activity (Fig. 1C and 2).[32] Interestingly, we found that Ib is a special subgroup containing only AtARFs, implying that these AtARFs were derived through a long-term evolution of *Arabidopsis* for *Arabidopsis*-specific functions (Fig. 2). It is worthy to notice that all the *AtARFs* of Group Ib are localized on only one chromosome (chromosome I). Group IV is a diverse group comprising GmARFs and AtARFs with variable MRs and CTDs. However, the GmARFs and AtARFs of Group IV have one common feature; they all contain the additional short segment of 12−44 residues within their DBD (Fig. 1C and 2).

Strong lines of evidence suggest that phylogenetic analysis enables functional prediction of various genes, including TF-encoding genes. For instance, phylogenetic analyses of the GmAP2_EREBP and GmNAC families of soybean and ONAC family of rice with their orthologs from other plant species, whose functions or stress-responsive expression patterns are known, resulted in a nearly perfect match between sequence conservation and functions or expression patterns.[67,68,49] Thus, phylogenetic-based functional prediction might quickly allow us to select candidate genes with positive functions in drought-stress responses from large gene families, which could be subsequently prioritized for further *in planta* functional studies. In *Arabidopsis*, mutations in the paralogous *AtARF01* and *AtARF02* resulted in delayed leaf senescence and floral organ abscission.[69,70] On the basis of our phylogenetic analysis (Fig. 2), the GmARF17, 27 and 44 and GmARF07, 12, 11 and 18, which are clustered with AtARF01 and AtARF02, respectively, might have similar functions to those of AtARF01 and AtARF02. These GmARFs might be selected as potential candidates for in-depth functional characterization with the aim to delay leaf senescence by genetic engineering, which in turn could enhance stress tolerance. Similarly, *AtARF07* and *AtARF19* were shown to play a positive role in regulation of lateral root development,[71] which is an important stress-related root trait for plant biotechnology.[5] Therefore, their closely homologous GmARF20, 31, 43 and 46 would gain a great attention of researchers who work to enhance this trait (Fig. 2).

3.3. Analysis of expression patterns of GmARF genes in different tissues and organs under well-watered conditions

In the next line of our study, we have interest in gaining knowledge about tissue-specific expression of

the *GmARF*s, because it enables us to identify the genes that are involved in defining the precise nature of individual tissues. Plants with an extensive fibrous root system and/or longer taproot can adapt better to drought stress, as they can forage subsoil surface moisture and/or reach lower soil layers where water is more readily available. On the other hand, plants with moderate shoot growth can survive longer water deficit conditions by minimizing evaporative leaf surface area and consuming less water. An appropriate control of plant architecture by genetic engineering is a promising approach for the development of crop varieties with enhanced drought tolerance and productivity.[5,23,72,73] Moreover, identification of tissue-specific genes, for instance root-specific genes, provides a resource of root-specific promoters for the improvement of drought tolerance by the enhancement of root growth.[20,74]

Thus, as a means to identify *GmARF* candidate genes that could be potentially used for enhancing drought tolerance by altering plant architecture, specifically shoot and/or root growth, when overexpressed or repressed in transgenic plant systems, we determined expression profiles for all 51 *GmARF* genes in the roots and shoots of 12-day-old soybean seedlings using qRT-PCR. We could detect the transcript of all *GmARF*

genes, whose expression levels were widely divergent. Based on their transcript abundance, 51 *GmARF* genes were classified into six groups (Fig. 3A−F), in which the highest/lowest expression ratios in roots (*GmARF05* versus *GmARF19*) and shoots (*GmARF12* versus *GmARF19*) were astonishingly huge with more than 490 309-fold and 165 905-fold, respectively. Using the criterion of 3-fold cut-off value, we found that, of 51 *GmARF*s, 11 genes were specifically expressed in roots, namely *GmARF05* (Fig. 3A), *GmARF09* (Fig. 3B), *GmARF02*, *18*, *22*, *27* and *49* (Fig. 3C), *GmARF15*, *28* and *33* (Fig. 3D) and *GmARF32* (Fig. 3E). Seven of the 51 *GmARF* genes displayed 3-fold higher expression in shoots and the remaining 33 genes showed ubiquitous expression patterns in both root and shoot tissues of young soybean seedlings. Specifically, the shoot-specific genes were grouped in Groups C (*GmARF35*), D (*GmARF25*, *29*, *34*, *36* and *48*) and E (*GmARF 50*) (Fig. 3). Additionally, *GmARF33* and *GmARF50* were found to be the most root- and shoot-specific genes. *GmARF33* more preferably expressed in roots than in shoots of 12-day-old soybean seedlings with the root/shoot ratio of ~436-fold, whereas *GmARF50* in shoots than in roots with the shoot/root ratio of about 100-fold.
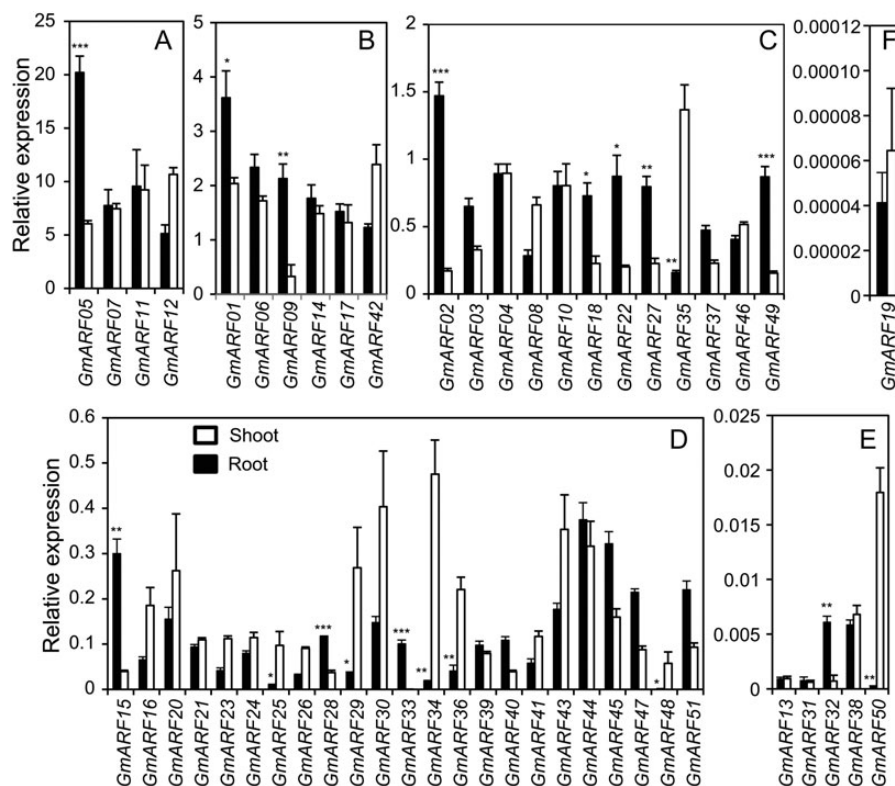

Figure 3. Expression patterns of 51 putative *GmARF* genes in roots (black bars) and shoots (white bars) of 12-day-old soybean seedlings under normal conditions. On the basis of their expression levels, the *GmARF* genes were classified into six groups (A−F). Data represent the means and standard errors of three independent biological samples. Asterisks indicate significant differences as determined by Student's *t*-test (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$). Relative expression was calculated based on the expression level of the target gene versus the level of the *60s* reference gene.

Recently, Libault *et al.*[59,60] reported a transcriptome atlas of soybean genes in eight tissues (nodules, roots, root hairs, root tips, leaves, flowers, green pods and SAM) using Illumina sequencing of soybean short transcripts. Thus, we also utilized these data to provide an overview about their expression patterns in these eight tissues. As shown in the heat map representation (Fig. 4), most of the *GmARF*s exhibited divergent expression profiles in the eight tissues examined. Based on their transcript abundance, the *GmARF*s could be classified into three major groups. Several genes displayed tissue-specific expression patterns; for instance, *GmARF02* and *05* that exhibited root organ-specific expression patterns (Fig. 4), which is consistent with our qRT-PCR analysis (Fig. 3A and C). A number of *GmARF*s showed their highest transcript abundance in SAM and/or green pods, such as *GmARF24, 34, 42, 19, 33* and *45*. Collectively, these observations demonstrate that the *GmARF*s have diverse expression patterns as their *Arabidopsis* counterparts,[32] suggesting that the functions of the GmARFs may be diversified in a similar manner as that of the AtARFs. It is worthy to mention that the duplicated gene pairs displayed similar expression profiles in the eight tissues examined although with different expression levels (Fig. 4). For instance, *GmARF45* has very high transcript abundance in SAM, but its expression is almost negligible in other seven tissues examined (Fig. 4). Similarly, *GmARF13*—the most closely homologue of *GmARF45* (Fig. 1B)—also specifically expressed in SAM. Other duplicated pairs, such as *GmARF02* and *05* and *GmARF11* and *18*, also displayed very similar expression profiles in the eight examined tissues, suggesting that these duplications were very likely resulted from the whole-genome duplication events.

With the progress in microarray analyses of soybean at the whole-genome-wide level using Affymetrix Genechips, a huge amount of data are also available for the evaluation of expression of soybean genes in various tissues. These data were collected by Genevestigator' developers, then analysed and housed on their database (https://www.genevestigator.com/).[58] Taking the advantage of the availability of these data, we expanded our expression study to examine the specific expression of *GmARF* genes in all 68 tissues and organs of soybean. This data set allowed us to analyse the expression of 42 of 51 *GmARF*s in total. The heat map shown in (Supplementary Fig. S2) displays the expression patterns of these *GmARF* genes, which may provide the temporal and spatial evidence linking them to their *in planta* functions. The expression data showed a high variability in transcript abundance of the *GmARF* genes in various tissues and organs, strongly indicating the diversified functions of the GmARF TFs in plant growth and developments. Expression of all 42 *GmARF* genes was very low or not observed at all in flower organs examined, such as
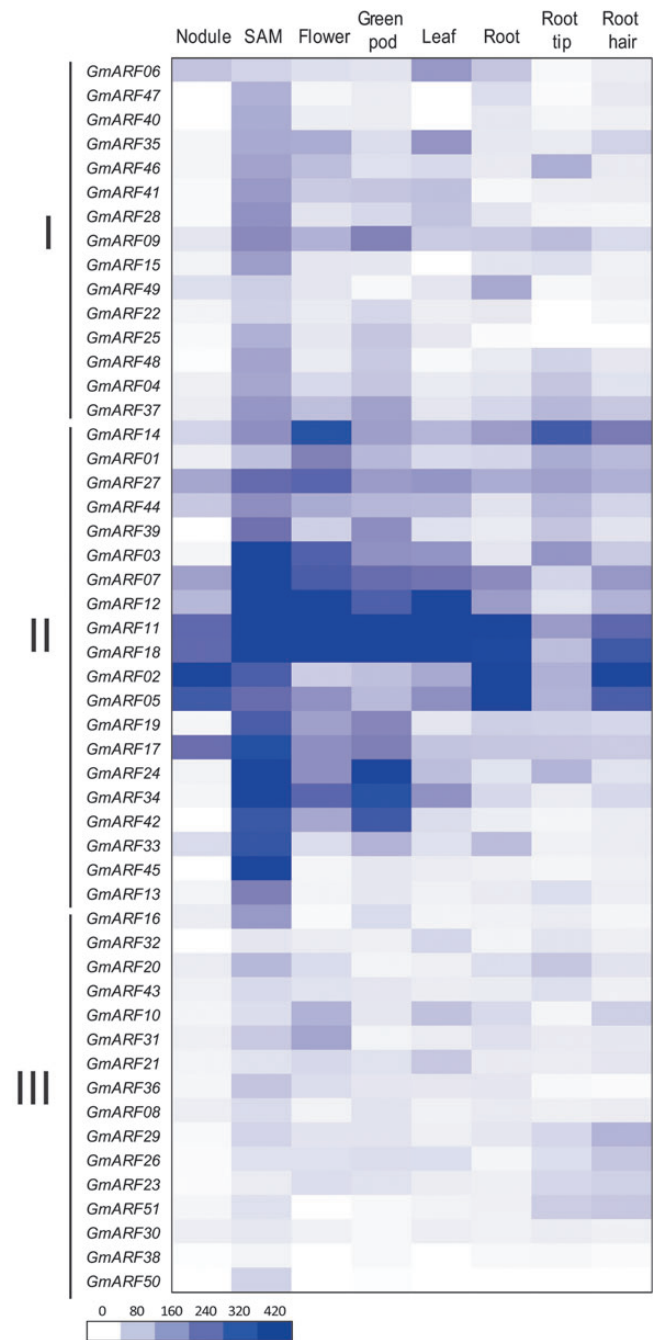

Figure 4. Heat map representation for tissue-specific expression of 51 *GmARF* genes in soybean. Expression patterns of the *GmARF* genes in eight indicated tissues were analysed using the Illumina transcriptome data. Elevated expression levels are indicated by increasing intensities of blue colour (saturated at 420) expressed in the normalized Illumina-Solexa read number.

pollen and stamen. Additionally, the transcripts of a few genes, such as *GmARF01, 06, 17, 46* and *49*, could be detected only in several organs among the 68 organs analysed.

The information obtained on tissue-specific expression of the *GmARF* genes can be used to address the combinatorial usage of GmARF TFs, allowing us to gain

an insight into the transcriptional programme of different tissues which is under the control of the GmARFs. Combinations of specific GmARFs with other type(s) of TFs might also regulate tissue-specific downstream genes. Protein−protein interactions, such as specific homodimerizations and heterodimerizations, modular flexibility and post-transcriptional and post-translational modifications, which are known to play important roles in determination of the functions of the TFs,[7,15,75−77] may also influence the functional specificity of the GmARFs. Analyses of these regulatory processes will enable us to elucidate the regulatory functions of the GmARF TFs in a comprehensive manner.

3.4. Analysis of expression patterns of the GmARF genes in roots and shoots during dehydration stress using qRT-PCR

With ∼4−7% of the genes encoding TFs in plant genome,[33,34,51,52,78] the TFs have been shown to play important roles in the regulation of environmental stress responses, including drought stress.[8,11,14,79] A growing body of evidence has demonstrated that auxin and the ARFs are implicated in drought-stress response,[25−30] suggesting that the stress-responsive ARF genes may be used to enhance drought tolerance in plants via genetic engineering. As a means to identify dehydration-responsive GmARF genes that are potentially used for in-depth characterization and engineering of soybean cultivars with improved drought tolerance, we performed a systematic expression profiling of the GmARF genes prior to launching laborious *in planta* functional studies. All 51 identified GmARFs were subjected to a comprehensive qRT-PCR analysis to assess their dehydration-responsive expression in root and shoot tissues of 12-day-old soybean plants that had been dehydrated for 2 and 10 h. The evaluations of expression patterns in roots and shoots separately, rather than in whole plants, might provide helpful information on the mode of action of stress-responsive GmARF genes in these individual tissues.

As shown in Figs 5 and 6, among 51 GmARFs 33 and 33 genes were found to be dehydration-responsive in shoots and/or roots of 12-day-old soybean seedlings. Specifically, with the criterion of 2-fold, a total of 25 and 8 GmARF genes were identified as up-regulated and down-regulated, respectively, in the shoots by dehydration (Fig. 5), whereas 5 and 28 genes as induced and repressed, respectively, in the roots by the same treatment (Fig. 6A and B). Additionally, GmARF33 and GmARF50 were the most induced genes by dehydration in shoots and roots, respectively, with >585- and 1320-fold inductions detected for GmARF33 and >15- and 30-fold increases in transcript levels observed for GmARF50, after 2 and 10 h of dehydration. Therefore, these two genes would be excellent

candidates for further *in planta* studies in soybean. A Venn diagram analysis indicated that two (GmARF12 and 50) of the up-regulated and seven (GmARF20, 26, 34, 35, 41, 43 and 51) of the down-regulated genes identified in roots and shoots were overlapped (Fig. 6C). On the other hand, of 30 GmARF genes that were down-regulated in roots, 12 genes (GmARF09, 10, 15, 18, 21, 27, 28, 33, 37, 38, 44 and 49) were found to be up-regulated in shoots (Fig. 6C). Expression levels of the GmARF genes that did not respond to dehydration were not shown.

3.5. Differential expression analysis of the GmARF genes in drought-stressed V6 and R2 soybean leaves and dehydrated shoots and roots of young soybean seedlings

As previously shown, dehydration stress altered the expression of many GmARF genes in roots and shoots of 12-day-old soybean seedlings. Recently, using the 66 K Affymetrix Soybean Array GeneChip, we have carried out genome-wide expression profiling of soybean leaves at V6 and R2 stages under drought stress.[61] This microarray data set allowed us to assess the drought-responsive expression patterns of the GmARF genes in the leaves of mature soybean plants. Among 51 GmARFs, three (GmARF12, 29 and 51) genes were found to be up-regulated and nine (GmARF03, 20, 23, 24, 25, 26, 28, 36 and 41) genes down-regulated by >2-fold (*q*-value <0.05) in drought-stressed V6 and/or R2 leaves (Supplementary Table S3).

Expression analysis of all 51 GmARFs in dehydrated shoots and roots of 12-day-old soybean seedlings using qRT-PCR has found 33 and 33 GmARF genes up-regulated or down-regulated in dehydrated shoot and/or root tissues (Figs 5 and 6). Comparative expression analysis of the GmARF genes in drought-stressed V6 and R2 leaves and dehydrated shoot and root tissues of 12-day-old soybean seedlings revealed that the majority of the GmARF genes exhibited highly variable responsiveness to water stress in the tissues examined (Supplementary Fig. S3 and Supplementary Table S3), indicating that the GmARF TFs may have specific functions in different tissues at different developmental stages under stress conditions. For instance, expression of GmARF51 was induced in drought-stressed V6 leaves but strongly repressed in dehydrated roots and shoots of young soybean seedlings, whereas that of GmARF18 was repressed in roots but induced in shoots of soybean seedlings by dehydration treatment and relatively unchanged in V6 and R2 leaves under drought stress (Supplementary Fig. S3). We also observed that, even in the same leaf tissue, the responsiveness of several GmARF genes, such as that of GmARF03, 20 and 51, to drought treatment was
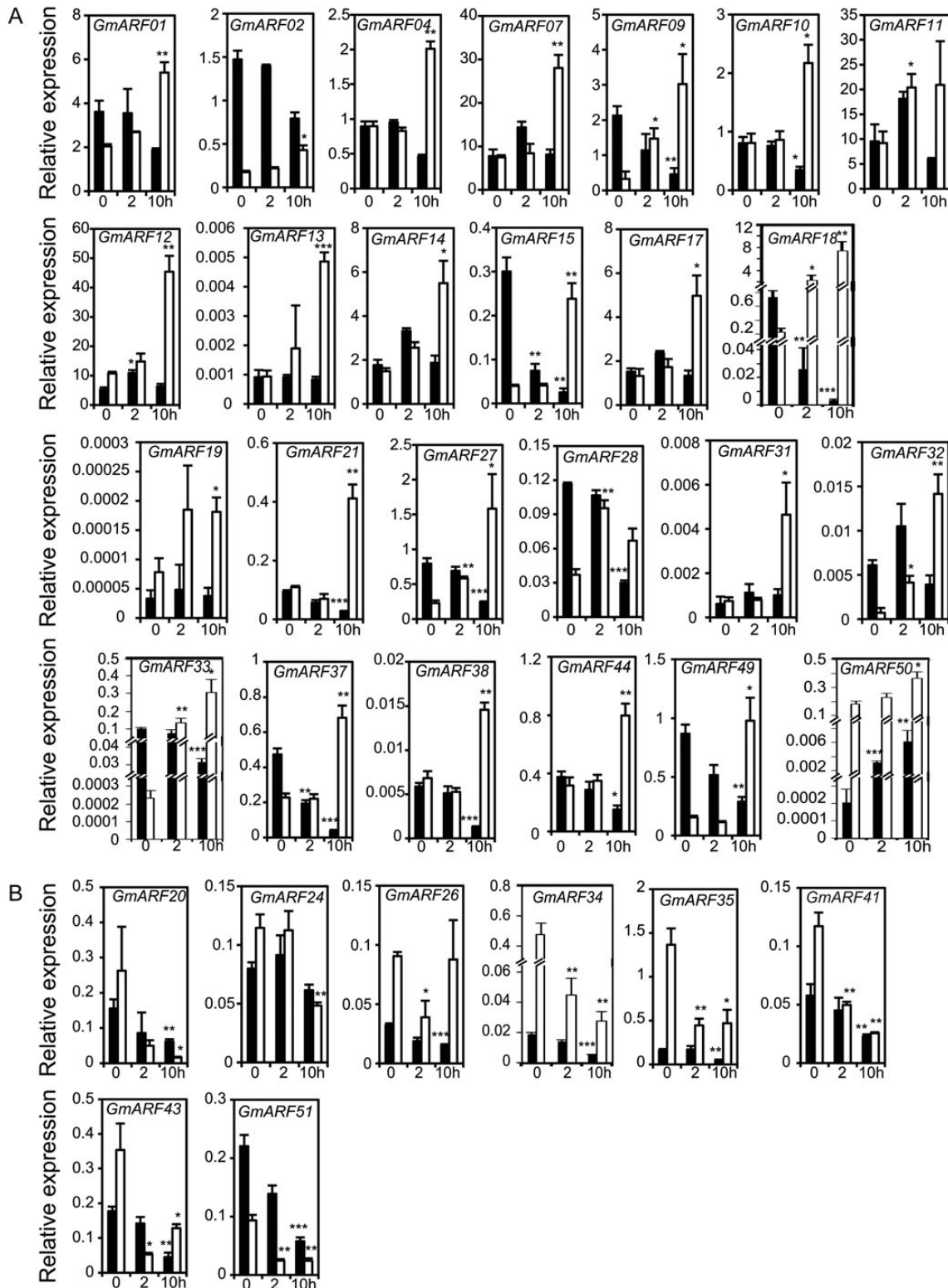
Figure 5. Expression of *GmARF* genes in roots (black bars) and shoots (white bars) of soybean plants under dehydration stress. (A) Up-regulated *GmARF* genes in shoots by at least 2-fold. (B) Down-regulated *GmARF* genes in shoots by at least 2-fold. Data represent the means and standard errors of three independent biological samples. Asterisks on the top of bars indicate significant differences as determined by Student's *t*-test (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$). Relative expression was calculated based on the expression level of the target gene versus the level of the *60s* reference gene.
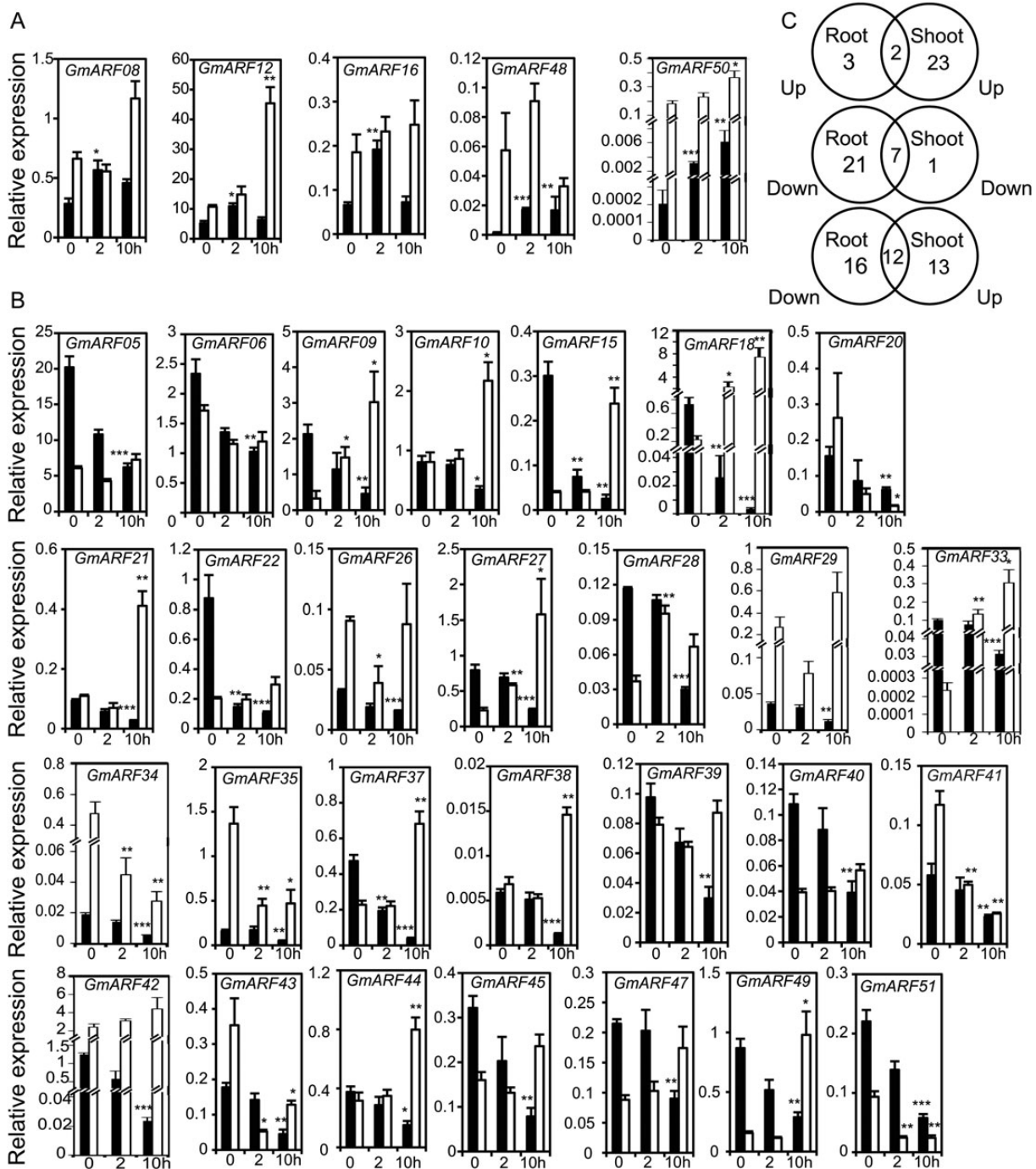
Figure 6. Expression of *GmARF* genes in roots (black bars) and shoots (white bars) of soybean plants under dehydration stress. (A) Up-regulated *GmARF* genes in roots by at least 2-fold. (B) Down-regulated *GmARF* genes in roots by at least 2-fold. (C) Venn diagram analysis of differentially expressed *GmARF* genes in shoots and roots of soybean seedlings. Data represent the means and standard errors of three independent biological samples. Asterisks on the top of bars indicate significant differences as determined by Student's *t*-test (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$). Relative expression was calculated based on the expression level of the target gene versus the level of the *60s* reference gene.

different at V6 and R2 stages. Collectively, this comparative analysis suggests that the dynamics of water stress-responsive expression of the *GmARF* genes in soybean is complex. Water stress may trigger different stress-responsive gene expression in different tissues at the same developmental stage or in the same tissue at different developmental stages.

3.6. Conclusion

The designed systematic characterization of the GmARF family in soybean has revealed key features in the structures of the GmARFs and in the relevant functions of this TF family in plant growth and development and drought-stress responses. The determination of expression patterns of the *GmARFs* genes in various tissues

and organs will enable us to identify those *ARF* genes that are expressed in limited specific region or in temporally regulated fashion. Studies of chromosomal distribution and duplications of the *GmARF* genes have provided valuable insights on the evolutionary aspects of soybean genome. Given that auxin is critical for organogenesis and embryo development, there is no doubt that the *GmARF* genes have immense and diverse roles in the life of soybean plants. The results of a comprehensive expression analysis of all the identified *GmARF* genes under normal and water stress conditions in different tissues and organs of soybean plants will help orient directions of molecular genetic studies, leading to better understanding of the functions of the GmARF TFs in soybean and their future applications. Overall, this study has enabled us to select water stress-responsive *GmARF* genes with more confidence for further *in planta* studies with the ultimate goal of development of improved drought-tolerant soybean cultivars by genetic engineering.

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

Funding

References

1. Sakai, T. and Kogiso, M. 2008, Soy isoflavones and immunity, *J. Med. Invest.*, **55**, 167–73.
2. Choudhary, S.P. and Tran, L.S. 2011, Phytosterols: perspectives in human nutrition and clinical therapy, *Curr. Med. Chem.*, **18**, 4557–67.
3. Burris, R.H. and Roberts, G.P. 1993, Biological nitrogen fixation, *Annu. Rev. Nutr.*, **13**, 317–35.
4. Sulieman, S. and Tran, L.S. 2012, Asparagine: an amide of particular distinction in the regulation of symbiotic nitrogen fixation of legumes, *Crit. Rev. Biotechnol.*, doi:10.3109/07388551.2012.695770.
5. Manavalan, L.P., Guttikonda, S.K., Tran, L.S. and Nguyen, H.T. 2009, Physiological and molecular approaches to improve drought resistance in soybean, *Plant Cell Physiol.*, **50**, 1260–76.
6. Tran, L.S. and Mochida, K. 2010, Functional genomics of soybean for improvement of productivity in adverse conditions, *Funct. Integr. Genomics*, **10**, 447–62.
7. Tran, L.S., Nakashima, K., Shinozaki, K. and Yamaguchi-Shinozaki, K. 2007, Plant gene networks in osmotic stress response: from genes to regulatory networks, *Methods Enzymol.*, **428**, 109–28.
8. Yamaguchi-Shinozaki, K. and Shinozaki, K. 2006, Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses, *Annu. Rev. Plant. Biol.*, **57**, 781–803.
9. Valliyodan, B. and Nguyen, H.T. 2006, Understanding regulatory networks and engineering for enhanced drought tolerance in plants, *Curr. Opin. Plant Biol.*, **9**, 189–95.
10. Hadiarto, T. and Tran, L.S. 2011, Progress studies of drought-responsive genes in rice, *Plant Cell Rep.*, **30**, 297–310.
11. Tran, L.S., Nishiyama, R., Yamaguchi-Shinozaki, K. and Shinozaki, K. 2010, Potential utilization of NAC transcription factors to enhance abiotic stress tolerance in plants by biotechnological approach, *GM Crops*, **1**, 32–9.
12. Yang, S., Vanderbeld, B., Wan, J. and Huang, Y. 2010, Narrowing down the targets: towards successful genetic engineering of drought-tolerant crops, *Mol. Plant.*, **3**, 469–90.
13. Puranik, S., Sahu, P.P., Srivastava, P.S. and Prasad, M. 2012, NAC proteins: regulation and role in stress tolerance, *Trends Plant Sci.*, **17**, 369–81.
14. Thao, N.P. and Tran, L.S. 2012, Potentials toward genetic engineering of drought-tolerant soybean, *Crit. Rev. Biotechnol.*, **32**, 349–62.
15. Jogaiah, S., Ramsandra Govind, S. and Tran, L.S. 2013, System biology-based approaches towards understanding drought tolerance in food crops, *Crit. Rev. Biotechnol.*, **33**, 23–39.
16. Finet, C., Fourquin, C., Vinauger, M., et al. 2010, Parallel structural evolution of auxin response factors in the angiosperms, *Plant J.*, **63**, 952–9.
17. Kieffer, M., Neve, J. and Kepinski, S. 2010, Defining auxin response contexts in plant development, *Curr. Opin. Plant Biol.*, **13**, 12–20.
18. de Jong, M., Wolters-Arts, M., Garcia-Martinez, J.L., Mariani, C. and Vriezen, W.H. 2011, The *Solanum lycopersicum* AUXIN RESPONSE FACTOR 7 (SlARF7) mediates cross-talk between auxin and gibberellin signalling during tomato fruit set and development, *J. Exp. Bot.*, **62**, 617–26.
19. Lau, S., De Smet, I., Kolb, M., Meinhardt, H. and Jurgens, G. 2011, Auxin triggers a genetic switch, *Nat. Cell Biol.*, **13**, 611–5.
20. Ha, S., Vankova, R., Yamaguchi-Shinozaki, K., Shinozaki, K. and Tran, L.S. 2012, Cytokinins: metabolism and function in plant adaptation to environmental stresses, *Trends Plant Sci.*, **17**, 172–9.
21. Choudhary, S.P., Yu, J.Q., Yamaguchi-Shinozaki, K., Shinozaki, K. and Tran, L.S. 2012, Benefits of brassinosteroid crosstalk, *Trends Plant Sci.*, **17**, 594–605.
22. Park, J.E., Park, J.Y., Kim, Y.S., et al. 2007, GH3-mediated auxin homeostasis links growth regulation with stress adaptation response in Arabidopsis, *J. Biol. Chem.*, **282**, 10036–46.
23. Zhang, S.W., Li, C.H., Cao, J., et al. 2009, Altered architecture and enhanced drought tolerance in rice via the

down-regulation of indole-3-acetic acid by TLD1/ OsGH3.13 activation, *Plant Physiol.*, **151**, 1889−901.

24. Zahir, Z.A., Shah, M.K., Naveed, M. and Akhter, M.J. 2010, Substrate-dependent auxin production by *Rhizobium phaseoli* improves the growth and yield of *Vigna radiata* L. under salt stress conditions, *J Microbiol. Biotechnol.*, **20**, 1288−94.

25. Lee, M., Jung, J.H., Han, D.Y., Seo, P.J., Park, W.J. and Park, C.M. 2012, Activation of a flavin monooxygenase gene YUCCA7 enhances drought resistance in *Arabidopsis*, *Planta*, **235**, 923−38.

26. Du, H., Wu, N., Fu, J., et al. 2012, A GH3 family member, OsGH3-2, modulates auxin and abscisic acid levels and differentially affects drought and cold tolerance in rice, *J. Exp. Bot.*, **63**, 6467−80.

27. Zhao, F.Y., Hu, F., Zhang, S.Y., Wang, K., Zhang, C.R. and Liu, T. 2013, MAPKs regulate root growth by influencing auxin signaling and cell cycle-related gene expression in cadmium-stressed rice, *Environ. Sci. Pollut. Res. Int.*, doi:10.1007/s11356-013-1559-3.

28. Song, Y., Wang, L. and Xiong, L. 2009, Comprehensive expression profiling analysis of OsIAA gene family in developmental processes and in response to phytohormone and stress treatments, *Planta*, **229**, 577−91.

29. Jain, M. and Khurana, J.P. 2009, Transcript profiling reveals diverse roles of auxin-responsive genes during reproductive development and abiotic stress in rice, *FEBS J.*, **276**, 3148−62.

30. Wang, S., Bai, Y., Shen, C., et al. 2010, Auxin-related gene families in abiotic stress response in *Sorghum bicolor*, *Funct. Integr. Genomics*, **10**, 533−46.

31. Shkolnik-Inbar, D. and Bar-Zvi, D. 2010, ABI4 mediates abscisic acid and cytokinin inhibition of lateral root formation by reducing polar auxin transport in Arabidopsis, *Plant Cell*, **22**, 3560−73.

32. Guilfoyle, T.J. and Hagen, G. 2007, Auxin response factors, *Curr. Opin. Plant Biol.*, **10**, 453−60.

33. Perez-Rodriguez, P., Riano-Pachon, D.M., Correa, L.G., Rensing, S.A., Kersten, B. and Mueller-Roeber, B. 2010, PlnTFDB: updated content and new features of the plant transcription factor database, *Nucleic Acids Res.*, **38**, D822−7.

34. Zhang, H., Jin, J., Tang, L., et al. 2011, PlantTFDB 2.0: update and improvement of the comprehensive plant transcription factor database, *Nucleic Acids Res.*, **39**, D1114−7.

35. Tiwari, S.B., Hagen, G. and Guilfoyle, T. 2003, The roles of auxin response factor domains in auxin-responsive transcription, *Plant Cell*, **15**, 533−43.

36. Hardtke, C.S., Ckurshumova, W., Vidaurre, D.P., et al. 2004, Overlapping and non-redundant functions of the Arabidopsis auxin response factors MONOPTEROS and NONPHOTOTROPIC HYPOCOTYL 4, *Development*, **131**, 1089−100.

37. Okushima, Y., Overvoorde, P.J., Arima, K., et al. 2005, Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in *Arabidopsis thaliana*: unique and overlapping functions of ARF7 and ARF19, *Plant Cell*, **17**, 444−63.

38. Xing, H., Pudake, R.N., Guo, G., et al. 2011, Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize, *BMC Genomics*, **12**, 178.

39. Wang, Y., Deng, D., Shi, Y., Miao, N., Bian, Y. and Yin, Z. 2012, Diversification, phylogeny and evolution of auxin response factor (ARF) family: insights gained from analyzing maize ARF genes, *Mol. Biol. Rep.*, **39**, 2401−15.

40. Shen, C., Wang, S., Bai, Y., et al. 2010, Functional analysis of the structural domain of ARF proteins in rice (*Oryza sativa* L.), *J. Exp. Bot.*, **61**, 3971−81.

41. Wu, J., Wang, F., Cheng, L., et al. 2011, Identification, isolation and expression analysis of auxin response factor (ARF) genes in *Solanum lycopersicum*, *Plant Cell Rep.*, **30**, 2059−73.

42. Mun, J.H., Yu, H.J., Shin, J.Y., Oh, M., Hwang, H.J. and Chung, H. 2012, Auxin response factor gene family in *Brassica rapa*: genomic organization, divergence, expression, and evolution, *Mol. Genet. Genomics*, **287**, 765−84.

43. Schmutz, J., Cannon, S.B., Schlueter, J., et al. 2010, Genome sequence of the palaeopolyploid soybean, *Nature*, **463**, 178−83.

44. Liao, Y., Zou, H.F., Wang, H.W., et al. 2008, Soybean GmMYB76, GmMYB92, and GmMYB177 genes confer stress tolerance in transgenic Arabidopsis plants, *Cell Res.*, **18**, 1047−60.

45. Tran, L.S., Quach, T.N., Guttikonda, S.K., et al. 2009, Molecular characterization of stress-inducible GmNAC genes in soybean, *Mol. Genet. Genomics*, **281**, 647−64.

46. Pinheiro, G.L., Marques, C.S., Costa, M.D., et al. 2009, Complete inventory of soybean NAC transcription factors: sequence conservation and expression analysis uncover their distinct roles in stress response, *Gene*, **444**, 10−23.

47. Mochida, K., Yoshida, T., Sakurai, T., Yamaguchi-Shinozaki, K., Shinozaki, K. and Tran, L.S. 2010, Genome-wide analysis of two-component systems and prediction of stress-responsive two-component system members in soybean, *DNA Res.*, **17**, 303−24.

48. Le, D.T., Nishiyama, R., Watanabe, Y., et al. 2011, Genome-wide expression profiling of soybean two-component system genes in soybean root and shoot tissues under dehydration stress, *DNA Res.*, **18**, 17−29.

49. Le, D.T., Nishiyama, R., Watanabe, Y., et al. 2011, Genome-wide survey and expression analysis of the plant-specific NAC transcription factor family in soybean during development and dehydration stress, *DNA Res.*, **18**, 263−76.

50. Ha, C.V., Le, D.T., Nishiyama, R., et al. 2013, Characterization of the newly developed soybean cultivar DT2008 in relation to the model variety W82 reveals a new genetic resource for comparative and functional genomics for improved drought tolerance, *BioMed. Res. Int.*, **2013**, 759657.

51. Mochida, K., Yoshida, T., Sakurai, T., Yamaguchi-Shinozaki, K., Shinozaki, K. and Tran, L.S. 2009, In silico analysis of transcription factor repertoire and prediction of stress responsive transcription factors in soybean, *DNA Res.*, **16**, 353−69.

52. Mochida, K., Yoshida, T., Sakurai, T., Yamaguchi-Shinozaki, K., Shinozaki, K. and Tran, L.S. 2010, LegumeTFDB: an integrative database of *Glycine max*,

    Lotus japonicus and *Medicago truncatula* transcription factors, *Bioinformatics*, **26**, 290−1.

53. Wang, Z., Libault, M., Joshi, T., et al. 2010, SoyDB: a knowledge database of soybean transcription factors, *BMC Plant Biol.*, **10**, 14.

54. Cheung, J., Estivill, X., Khaja, R., et al. 2003, Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence, *Genome. Biol.*, **4**, R25.

55. Tang, H., Bowers, J.E., Wang, X., Ming, R., Alam, M. and Paterson, A.H. 2008, Synteny and collinearity in plant genomes, *Science*, **320**, 486−8.

56. Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. 1997, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res.*, **25**, 4876−82.

57. Tamura, K., Dudley, J., Nei, M. and Kumar, S. 2007, MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0, *Mol. Biol. Evol.*, **24**, 1596−9.

58. Hruz, T., Laule, O., Szabo, G., et al. 2008, Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes, *Adv. Bioinformatics*, **2008**, 420747.

59. Libault, M., Farmer, A., Brechenmacher, L., et al. 2010, Complete transcriptome of the soybean root hair cell, a single-cell model, and its alteration in response to *Bradyrhizobium japonicum* infection, *Plant Physiol.*, **152**, 541−52.

60. Libault, M., Farmer, A., Joshi, T., et al. 2010, An integrated transcriptome atlas of the crop model *Glycine max*, and its use in comparative analyses in plants, *Plant J.*, **63**, 86−99.

61. Le, D.T., Nishiyama, R., Watanabe, Y., et al. 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.

62. Le, D.T., Aldrich, D.L., Valliyodan, B., et al. 2012, Evaluation of candidate reference genes for normalization of quantitative RT-PCR in soybean tissues under various abiotic stress conditions, *PLoS One*, **7**, e46487.

63. Zhang, J.Z. 2003, Evolution by gene duplication: an update, *Trends Ecol. Evol.*, **18**, 292−8.

64. Shoemaker, R.C., Schlueter, J. and Doyle, J.J. 2006, Paleopolyploidy and gene duplication in soybean and other legumes, *Curr. Opin. Plant Biol.*, **9**, 104−9.

65. Schlueter, J.A., Lin, J.Y., Schlueter, S.D., et al. 2007, Gene duplication and paleopolyploidy in soybean and the implications for whole genome sequencing, *BMC Genomics*, **8**, 330.

66. Ulmasov, T., Hagen, G. and Guilfoyle, T.J. 1997, ARF1, a transcription factor that binds to auxin response elements, *Science*, **276**, 1865−8.

67. Fang, Y., You, J., Xie, K., Xie, W. and Xiong, L. 2008, Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice, *Mol. Genet. Genomics*, **280**, 547−63.

68. Zhang, G., Chen, M., Chen, X., et al. 2008, Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L.), *J. Exp. Bot.*, **59**, 4095−107.

69. Ellis, C.M., Nagpal, P., Young, J.C., Hagen, G., Guilfoyle, T.J. and Reed, J.W. 2005, AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis thaliana*, *Development*, **132**, 4563−74.

70. Lim, P.O., Lee, I.C., Kim, J., et al. 2010, Auxin response factor 2 (ARF2) plays a major role in regulating auxin-mediated leaf longevity, *J. Exp. Bot.*, **61**, 1419−30.

71. Fukaki, H., Taniguchi, N. and Tasaka, M. 2006, PICKLE is required for SOLITARY-ROOT/IAA14-mediated repression of ARF7 and ARF19 activity during Arabidopsis lateral root initiation, *Plant J.*, **48**, 380−9.

72. Achard, P., Cheng, H., De Grauwe, L., et al. 2006, Integration of plant responses to environmentally activated phytohormonal signals, *Science*, **311**, 91−4.

73. Huang, J.G., Yang, M., Liu, P., Yang, G.D., Wu, C.A. and Zheng, C.C. 2009, GhDREB1 enhances abiotic stress tolerance, delays GA-mediated development and represses cytokinin signalling in transgenic Arabidopsis, *Plant Cell Environ.*, **32**, 1132−45.

74. Werner, T., Nehnevajova, E., Kollmer, I., et al. 2010, Root-specific reduction of cytokinin causes enhanced root growth, drought tolerance, and leaf mineral enrichment in Arabidopsis and tobacco, *Plant Cell*, **22**, 3905−20.

75. Tran, L.S., Nakashima, K., Sakuma, Y., et al. 2007, Co-expression of the stress-inducible zinc finger homeodomain ZFHD1 and NAC transcription factors enhances expression of the ERD1 gene in Arabidopsis, *Plant J.*, **49**, 46−63.

76. Qin, F., Sakuma, Y., Tran, L.S., et al. 2008, Arabidopsis DREB2A-interacting proteins function as RING E3 ligases and negatively regulate plant drought stress-responsive gene expression, *Plant Cell*, **20**, 1693−707.

77. Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A. and Luscombe, N.M. 2009, A census of human transcription factors: function, expression and evolution, *Nat. Rev. Genet.*, **10**, 252−63.

78. Mochida, K., Yoshida, T., Sakurai, T., Yamaguchi-Shinozaki, K., Shinozaki, K. and Tran, L.S. 2013, TreeTFDB: an integrative database of the transcription factors from six economically important tree crops for functional predictions and comparative and functional genomics, *DNA Res.*, **20**, 151−62.

79. Ha, S. and Tran, L.S. 2013, Understanding plant responses to phosphorus starvation for improvement of plant tolerance to phosphorus deficiency by biotechnological approaches, *Crit. Rev. Biotechnol.*, doi:10.3109/07388551.2013.783549.