Comprehensive RNA-seq Analysis Revealed Molecular Pathways and Genes Associated with Drought Tolerance in Wild Soybean (Glycine soja Sieb. & Zucc.)

Muqadas Aleem[1], Muhammad Raza[1], Muhammad Haider[2], Rana Atif[3], Zulfiqar Ali[4], Javaid Bhat[2], and Tuanjie Zhao[2]

[1]Nanjing Agricultural University College of Agriculture
[2]Nanjing Agricultural University
[3]University of Agriculture Faisalabad
[4]MNS-University of Agriculture

April 28, 2020

Abstract

Drought stress at germination stage is an important environmental stress limiting crop yield. Hence, our study investigated comparative root transcriptome profiles of four contrasting soybean genotypes viz., drought-tolerant (PI342618B/DTP & A214/DTL) and drought-sensitive (NN86-4/DSP & A195/DSL) under drought stress using RNA-Seq approach. Total of 4850 and 6272 differentially expressed genes (DEGs) were identified in tolerant (DTP & DTL) and sensitive (DSP & DSL) genotypes, respectively. Principle component analysis (PCA) and correlation analysis revealed higher correlation of DTP with DTL. Both gene ontology (GO) and MapMan analyses showed drought response was enriched in the DEGs associated with water and auxin transport, cell wall/membrane, antioxidant activity, catalytic activity, secondary metabolism, signaling and transcription factor (TF) activities. Out of 981 DEGs screened from above terms, only 547 showed consistent opposite expression between contrasting genotypes. Twenty-eight DEGs of 547 were located on Chr.08 rich in QTLs and "Hotspot regions" associated with drought stress, and eight of them showed non-synonymous SNP polymorphism. Hence, ten genes (including above eight genes plus two hub genes) were predicated as possible candidates regulating drought tolerance, which needs further functional validation. Overall, the transcriptome profiling provided in-depth understanding about the genetic mechanism and candidate genes underlying drought tolerance in soybean.

Introduction

Soybean is an economically important legume-cum-oilseed crop being rich source of protein (40%) and edible oil (20%) (Chaudhary *et al.* , 2015). However, soybean production is affected by a range of environmental stresses (Sharmin *et al.* , 2020). In recent years, increased occurrence of drought events was reported due to global climate change (Bailey-Serres et al., 2012). Drought stress at germination stage is one of the major constraints that badly affects seed germination and seedling establishment. Accumulated evidence indicates that yield is affected by drought stress at early (germination and vegetative) and late (reproductive) growth stages (Zhao *et al* ., 2017; Wijewardana *et al.* , 2019). However, a little progress has been made to elucidate the genetic mechanisms of drought tolerance during germination stage in soybean. Hence, concerted efforts are needed to understand the genetic basis and genes associated with drought tolerance during germination stage.

During germination stage many essential physiological and biochemical processes are induced to initiate germination such as hydrolysis, subcellular structures, cell elongation biosynthesis of macromolecules and

respiration (Bewley *et al* ., 2013). Many phytohormones i.e., ethylene (ETH), auxin (AUX), abscisic acid (ABA), gibberellic acid (GA), salicylic acid (SA) and jasmonic acid (JA) as well as some environmental conditions, including light, oxygen and temperature regulates seed germination (Miransari and Smith, 2014). Drought stress is an important environmental factor hampering seed germination and seedling vigor (Delachiave *et al.* , 2003; Ahmad *et al.* , 2009). It not only hinder the seed germination but also increases mean germination time (Valliyodan and Nguyen, 2006). At germination, soybean seed must absorb water (50% of its weight) to begin normal germination. Insufficient water availability especially during seed emergence stage lower the soybean yield (Wijewardana *et al.* , 2019; Kobraee *et al.* , 2011). Several researchers have reported the negative impact of drought on seed germination and seedling vigor in different crops such as mung bean (Rani *et al.* , 2018), barnyard grass (Wu *et al.* , 2019), sesame (Boureima *et al* . 2011), maize (Ahmad *et al.* , 2018) and soybean (Vieira *et al* . 1991).

Response of crop plants to drought stress at molecular level involve perception, signal transduction, gene expression and ultimately metabolic changes resulting in stress tolerance (Huang *et al.* , 2012). Drought tolerance being a complex quantitative trait involves numerous regulatory and functional genes (Jiménez *et al.* , 2013). Hence, numerous genes are involved in stomatal movement, osmolyte metabolism, antioxidant activity and phytohormones signalling, under drought stress and re-watering (Shinozaki and Yamaguchi-Shinozaki, 2007; Laxa *et al.* , 2019). Besides, drought stress stimulates many adaptive signaling pathways to cope with adverse effects of stress, such as MAPK cascade (Sinha *et al.* , 2011) and Ca^{2+}signaling (Wilkins *et al.* , 2016) and signaling of phytohormones (Tiwari *et al.* , 2017). These stress signaling pathways modify the expression of drought responsive genes and provide defensive mechanisms through activating downstream TFs, (Lan *et al.* , 2017) However, most of these gene function as well as regulatory networks involved in drought stress at germination stage have remained elusive. Hence, it is prerequisite to identify the genetic mechanism and networks involved in modulation of drought tolerance.

Advances in next-generation sequencing (NGS) platforms i.e., Illumina/Solexa has made it possible to understand the complexity and regulation of gene expression networks in different crop species under dehydration stress (Chen *et al.* , 2016). The NGS platforms provides wider and more robust transcriptome analyses (Severin *et al.* , 2010). Transcriptome studies (with Affymetrix/RNA sequencing technology) utilizing contrasting lines to understand drought stress response has been carried in different crops such as rice (Degenkolbe*et al* . 2009; Lenka *et al.* , 2011), but in case of soybean mostly single genotype was used (Le et al. 2012; Song *et al.* , 2016). Use of contrasting genotypes in transcriptome studies will provide comparative and detailed information on DEGs, that could possibly assist in identifying important genes regulating drought tolerance. Comparative transcriptomic analysis of various genotypes will be helpful to dissect biological pathways and mechanisms imparting tolerance to environmental stresses, as well as to explore genes under selection. Although, significant insights have been provided by fundamental research into the physiological and molecular responses of plants to drought stress, but the divergence in root transcriptome of soybean drought responsive genotypes during germination remains largely unexplored.

In the present study, the comparative transcriptomic analysis between drought-tolerant (PI342618B/DTP & A214/DTL) and drought-sensitive (NN86-4/DSP & A195/DSL) soybean genotypes were analyzed comprehensively to gain insights into soybean defense response under drought stress at germination stage. This study provided an in-depth information at the global transcriptome level, and identified pathways and candidate genes associated with drought tolerance at germination stage in soybean.

Material and Method:

Plant Material

Four soybean genotypes viz., PI342618B-drought-tolerant parent (DTP;*Glycine soja* Sieb. & Zucc.), NN86-4-drought-sensitive parent (DSP; *Glycine max* L.), A214-drought-tolerant line (DTL) and A195-drought-sensitive line (DSL) were used for investigation of drought stress response in the present study. The A214/DTL and A195/DSL are two introgression lines derived from NN86-4 × PI342618B cross through repeated backcrossing of three times with the cultivated recurrent parent NN86-4. All the plant materials were received

from National Center for Soybean Improvement, Nanjing Agricultural University, Jiangsu Province, China.

Growth Condition and Phenotypic Evaluation

Healthy and good quality seeds of all genotypes were used for phenotypic studies and other investigations. These seeds were first disinfected by using 70% (v/v) ethanol, and then surface-sterilized for 4h in a tightly sealed chamber using chlorine gas (100 ml NaClO + 15 ml HCl). To determine the optimum PEG 6000 concentration for drought-stress treatment, we tested different PEG concentrations viz., 10%, 15% & 20% (w/v). About 50-70 seeds of each genotype were placed in petri dishes (diameter is 15 cm) containing two filter paper at the bottom, and were filled with 20 mL of PEG 6000 solution.

Seed germination was evaluated under 15% PEG (w/v) and control (distilled water) conditions, respectively. Completely randomized design (CRD) were used for experiment trial with three biological replicates. Petri-plates containing sterilized seeds were incubated in a growth chamber for five days at temperature of 25°C and relative humidity (RH) of 60% under a 16 h/8 h (light/dark) cycle. Seeds with young radicle length approximately ˜1–2mm were considered as germinated seed. Seeds germination was observed every day for the estimation of the total germination percentage (GP). After four days, root length (RL) of each genotype was calculated under drought and control conditions. Primary roots from each genotype under both conditions were collected, and samples were named as DTL-C, DTL-T, DSL-C, DSL-T, DSP-C, DSP-T, DTP-C and DTP-T. From each genotype two biological replicates were used, and root tissue were quickly frozen in liquid nitrogen and stored at -80degC for RNA extraction. Following formula was used to calculate Germination rate (GR) :

GR= (germination rate under drought stress)/ (germination rate in control) x100%

Library Construction and RNA Sequencing

Total RNAs was extracted from the root samples of four soybean genotypes under control and stress conditions using TRIzol reagent by following the manufacturer's protocols (Invitrogen, Carlsbad, CA, USA). The RNA quantity and integrity was checked using Nano-Drop Spectrometer (ND-1000 Spectrophotometer, Thermo Scientific) and Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, US), respectively. To select the best quality RNA an integrity value [?]7 was used for library construction. Total RNA of 2 μg from each sample was used for the construction of paired-end libraries using Illumina TruSeq RNA Sample Preparation Kit (Illumina, San Diego, CA, USA), in accordance with the manufacturer's instructions. Qubit® 2.0Fluorometer (Life Technologies, USA) was used to quantify purified libraries, and are authenticated by Agilent 2100-bioanalyzer (Agilent Technologies, USA) to validate the fragment size and concentration. Furthermore, libraries were diluted up to 10 Pm for cBot clustering and RNA-Seq was performed using Illumina HiSeq 2500 platform by Shanghai Biotechnology Corporation.

Quality Filtering and Read Mapping

Short-fragment reads, rRNA reads, adapter/ primer and low quality reads were filtered from Fastq files of raw data by using Trimmomatic v 0.35 (Bolger *et al.* , 2014) and FASTX-toolkit present in the FastQC (http://www.bioinformatics. babraham.ac.uk/projects/fastqc/). High-quality filtered reads were mapped to *Glycine max* reference genome (*G max1.1 version*) available at Phytozome v9.0 database (Schmutz *et al.* , 2010; Goodstein *et al.* , 2012) using HISAT2 software (Kim et al., 2015).

Sequence Data Analysis

Only uniquely mapped reads were used for further analysis. Transcriptional level (FPKM) of all the samples were measured using Cufflinks software (Trapnell et al., 2012), while differentially expressed genes (DEGs) between different samples were estimated using Cuffdiff (Trapnell et al. 2012). Transcripts with a log2FC (fold change) [?] +2 and [?] -2, and FDR (False Discovery Rate) value <0.05 were selected as significant differentially expressed genes (DEGs).

Functional Classification and Pathway Identification

Principal component analysis (PCA) was performed to assess variability for the soybean genotypes under control and drought conditions using the PRCOMP command with default setting in the R software package (Robinson*et al.* , 2010). Log_2-transformed FPKM values of the DEGs were used for K-means clustering using Pearson correlation in Microarray Experiment Viewer (MeV, v4.9) software. AgriGO v2.0 database (http://systemsbiology.cau.edu.cn/agriGOv2/) was implemented for GO enrichment analysis of DEGs by using *Glycine max* as reference background, and DEGs were classified into three major categories: biological processes (BP), cellular components (CC) and molecular function (MF) (Tian *et al.* , 2017). Reduce and Visualize GO analyses /REVIGO (http://revigo.irb.hr/) were performed to remove the redundancy of GO terms using SimRel semantic similarity measure, with an allowed similarity of 0.7 (medium), and the results were displayed as scatter plot (Supek *et al.* , 2011). Up- and down-regulated transcripts were subjected to MapMan software version 3.6.0 RC1 (http://mapman.gabipd.org/web/guest/mapman) (Usadel *et al.* , 2009). Mapped gene intensity of fold change of various pathways (both biological or metabolic) were plotted by blue and red schema. The KEGG (Kyoto encyclopedia of gene and genome) was further utilized for pathway enrichment analysis of DEGs (Kanehisa *et al.* , 2008).

PPI Network and Cluster Analysis

STRING database (Search Tool for Retrieval of Interacting Genes) was used to construct the protein-protein interaction (PPI) network of Transcription factor (TFs) using *G. max* as reference (Szklarczyk*et al.* , 2014).

In order to predict candidate gene, differentially expressed genes (DEGs) screened were clustered using dChip software (Lin *et al* . 2004), and HeatMaps were also drawn using the same dChip software.

DNA Sequencing and Polymorphism Detection

Healthy and young leaves were taken from soybean seedling (seven-days old) grown in controlled conditions, and were stored in liquid nitrogen. Cetyltrimethylammonium bromide (CTAB) method was used for isolation of DNA from DTL/A214 and DSL/A195 genotypes (Murray and Thompson, 1980). Sequencing libraries (paired-end) having fragment size of ˜ 350 bp were developed according to the instructions of manufacturer's (Illumina Inc.), and Illumina HiSeq 2000 sequencer was used for sequencing of library by Novogene company. Low-quality and duplicate reads were filtered from raw data. Furthermore, BWA V0.7.8 software was used for mapping of processed reads to the*Glyma.Wm82. a2. v1* / reference genome (Li and Durbin, 2009). Polymorphism sites (SNPs and InDels) were detected by SAM tools (Stanke and Waack, 2003), and possible synonymous/non-synonymous SNP variation annotations was analyzed using SnpEFF software (Cingolani *et al* ., 2012).

Expression Validation by qRT-PCR

To validate and check the reliability of RNA-Seq data, qRT-PCR analysis was performed. Twelve stress induced genes were selected for experimental validation by qRT-PCR. From each biological replicate, cDNA was synthesized using the HiScriptII-QRT-superMix through qPCR+ gDNA wiper (Vazyme, Nanjing, China) in accordance with manufacturer's instruction. ABI 7500 Real-Time PCR System (Applied Biosystems, USA) was used for qRT-PCR analysis. Each reaction comprises of 10µl 2×SYBR Green Master Mix (Applied Biosystems, USA), 2.0µl cDNA, and 400 nM primer (forward, reverse) upto final volume of 20µl. The PCR conditions were 95@C for 3 min followed by 40 cycles of 95@C for 10 s, 60@C for 30 s and 65@C for 5s. Moreover, expression level were normalized by using actin gene (*GmActin11* , as an internal control) and Ct method (Livak and Schmittgen, 2001). For each sample, qRT-PCR reaction was repeated in two replicates, and three technical repeats of each biological replicate. The PRIMER5.0 software was used to design primers, and all primers are listed in Supplementary Table S1.

RESULTS:

Phenotypic Evaluation

To determine optimum PEG 6000 concentration for drought stress treatment, all the four soybean genotypes viz., DTP, DSP, DTL & DSL were subjected to different concentrations of PEG treatments viz., 10%,15% &

20% (w/v) in three biological replicates at germination stage, and the results revealed that 15% PEG is the optimum concentration for the evaluation of drought stress tolerance**(Fig 1A).** In the present study, GR and RL were used to evaluate drought tolerance of DTP, DSP, DTL & DSL genotypes under 15 % PEG and control conditions **(Fig 1B & C)** . The results revealed that all four genotypes showed normal germination under control conditions (**Fig 1B**); however, under stress conditions tolerant genotypes viz., DTP & DTL showed normal germination (GR=100%) at the 4^{th} day of stress treatment, whereas sensitive genotypes (DSP & DSL) revealed only 20% GR at the same time interval**(Fig. 1C).** Under control conditions, all genotypes showed normal RL, whereas the RL of all genotypes were reduced under stress relative to control condition (**Fig. 1D**). Furthermore, the RL of sensitive genotypes (DSP & DSL) were significantly (*P<0.01*) reduced under stress compared to tolerant genotypes viz., DTP & DTS **(Fig. 1D)** , and transgression effect has been identified in DTL for RL**.** Hence, our results revealed that DTP and DTL are tolerant to drought stress, whereas DSP and DSL are drought-sensitive, and suggested that these contrasting genotypes can be effectively used for the analysis of drought tolerance.

RNA-seq and Assembly

In total, the 493.15 million reads were produced from the transcriptome sequencing (RNA-seq) of eight samples that includes one sample from each four genotypes under stress and control conditions, with about 58.44 to 65.38 million reads per sample **(Table 1).** Reads with adapter contamination and low quality base ([?]Q20) were trimmed using Trimmomatic (Bolger *et al.* , 2014) and Fast-QCbox (Katta et al. 2015). As a result of this screening, we obtained a total of 465.80 million (˜94%) high quality reads, and total of 440.3 million reads (94.55%) were mapped on to soybean reference genome (v1.0) (**Table 1;** Schmutz *et al.* , 2010)**.**

Identification of DEGs Under Drought Stress Treatment

RNA-seq was used to detect DEGs in the soybean roots of DTP, DSP, DTL & DSL genotypes subjected to drought stress at germination stage. By comparing the genes showing differential expression under drought stress, a total of 2325, 3060, 2525 and 3212 DEGs (including both up- and down-) were found in the pairwise comparisons of DTP, DSP, DTL and DSL, respectively under drought treatment vs. control by using log2FC[?]2 and *q* value [?] 0.05 **(Fig. 2A).** Hence, Sensitive genotypes viz., DSP and DSL showed higher number of DEGs as compared to tolerant genotypes DTP & DTL **(Fig. 2A).** Moreover, in sensitive genotypes there are more number of down-regulated genes compared to tolerant genotypes. By comparing the transcriptome profiles of four genotypes, it was observed that total of 676, 493, 510 and 700 DEGs were uniquely and uniformly expressed in DTP, DSP, DTL and DSL, respectively under drought stress treatment (**Fig. 2B**). In addition, 775 DEGs were common among all four genotypes.

Further, we analyzed overlapping DEGs among 12 different combinations of four genotypes under drought and control conditions using Circos software **(Fig. 2C).** This analysis indicates that a major fraction of DEGs were unique under stress condition compared to control in all 12 combinations. Highest number of common DEGs (1950) among all genotype combination were found between tolerant genotypes combination under stress and control conditions (DTP-T vs DTL-T & DTP-C vs DTL-C), whereas it was only 550 between sensitive genotypes under stress and control conditions (DSP-T vs DSL-T & DSP-C vs DSL-C). In addition, higher number of unique DEGs (2847) were found in sensitive genotypes under drought stress (DSP-T vs DSL-T) compared to only 2095 in tolerant genotypes (DTP-T vs DTL-T). These findings are in agreement with DEG data and venn-diagram analysis, and revealed that sensitive genotypes are more influenced by drought stress compared to tolerant genotypes at the transcriptomic/gene expression level.

PCA and Correlation Analysis of DEGs

To reduced data dimensionality and visualize genotype relationship under stress and control conditions, we performed PCA analysis. The first principal component (PC1) displayed 30.7% of total variation, while as second principal component (PC2) explained 21.6% variation across the data set **(Fig. 3A).** Scores plot between the PC1 and PC2 shows clear separation of genotypes by PC1 under drought and control conditions. However, based on the gene expression variation, wild-parent (DTP/P1342618B) is clearly separated by PC2

from other three genotypes viz., DSP, DTL & DSL; however, among these three genotypes tolerant line (DTL) is more nearer to DTP/P1342618B genotype **(Fig. 3A)** . This provides some explanation about the differences in the gene expression pattern under drought and control conditions as well as in the tolerant and sensitive genotypes.

To validate the results of PCA analysis, we performed correlation analysis on the normalized expression values from all the samples, and created a dendrogram **(Fig. 3B).** In roots, all four control samples (DTP-C, DSP-C, DTL-C and DSL-C) were grouped in one cluster (cluster-I), whereas the four stress samples (DTP-T, DSP-T, DTL-T and DSL-T) were grouped in another cluster (cluster-II). In both clusters (I & II), wild-tolerant parent was grouped alone in separate sub-cluster, while other three genotypes are grouped together in another sub-cluster. This can be explained by the fact as both introgression lines viz., DTL and DSL are derived through repeated backcrossing of three times with recurrent parent NN86-4/DSP, and after three repeated backcross 93.8% of the recurrent parent genome is recovered in the introgression lines (Collard et al. 2005), and hence due to this genome similarity both introgression lines (DTL & DSL) are grouped together with recurrent parent NN86-4/DSP in the same sub-cluster. Among these three genotypes (DSP, DTL & DSL), the tolerant line DTL showed higher correlation with wild-tolerant parent DTP in both Cluster I & II under both conditions; however, under stress condition (Cluster II) correlation was relatively higher compared to control conditions. This indicates that genes for drought tolerance in DTL line are derived from DTP wild-parent. Hence, the results of correlation analysis are in agreement with that of PCA analysis **(Fig. 3A & B**). These results revealed that DTP is genetically more similar to DTL-tolerant line relative to DSL and DSP genotypes.

GO Classification and Enrichment Analysis of DEGs.

In order to evaluate the observed differences in gene expression profiles toward specific functions, we performed gene ontology (GO) analysis using Singular Enrichment Analysis (SEA) tool in agriGO using *Glycine max* as reference. A total of 2525 (1001 up-regulated and 1524 down-regulated), 3212 (1030 up-regulated and 2182 down-regulated), 3060 (860 up-regulated and 2200 down-regulated) and 2325 DEGs (496 up-regulated, and 1829 down-regulated) found in DTL, DSL, DSP & DTP, respectively were grouped in one of the three classes: biological processes (BP), cellular components (CC), molecular function (MF)**(Fig 4).** Moreover, the DEGs with unknown functional annotation (uncharacterized genes) can be considered as novel genes as they were not associated with any GO category **(Table S6).** Among the DEGs identified in DTL, 106, 24 and 64 were associated with BP, CC and MF terms, respectively **(Fig. 4A and Table S2**), while 107, 27 and 80 DEGs were identified with BP, CC and MF, respectively in case of DSL**(Fig. 4B Table S3**). However, in DTP total of 116 (BP), 71 (MF) and 34 (CC) GO terms were found, and for DSP 94, 76 and 40 GO terms were found for BP, MF and CC, respectively **(Fig. 4C & 4D; Table S4 & S5**). In all four soybean genotypes, a large number of DEGs were related with the terms viz., metabolic process, cellular process, signaling-organism process, catalytic activity, cell, cell part, membrane, membrane part, and binding **(Fig. 4).** These result suggests vital role of these GO terms in drought tolerance at germination stage in soybean.

To visualize difference of specific GO terms between the DTP vs DSP and DTL vs DSL genotypes, we used up- and down-regulated enriched GO categories (FDR < 0.05) from DTP vs DSP and DTL vs DSL to REVIGO analysis (**Table S2-S5**). GO terms associated with constantly expressed DEGs of DTP vs DSP and DTL vs DSL genotypes for BP, MF, and CC were shown in **Fig. 5 & 6** . Scatter plots displayed highly significant GO terms of BP i.e., water transport (GO:0006833), carbohydrate transport (GO:0008643), response to oxidative stress (GO:0006979), hormone-mediated signaling pathway (GO:0009755), Ca^{+}-mediated signaling (GO:0019722), secondary metabolite synthesis (GO:0044550), chaperone-mediated protein folding (GO:0061077), defense response (GO:0006952) and auxin transport (GO:0060918) were upregulated in tolerant genotypes (DTP & DTL)**(Fig. 5A & 6A),** whereas reverse trend was observed in drought-sensitive genotypes viz., DSP and DTL **(Fig. 5A & 6A**). However, significant GO term viz., signal transduction (GO:0007165) was uniquely found in tolerant genotypes. In case of MF, GO terms such as GO:0016209 (antioxidant activity), GO:0009055 (electron carrier activity), GO:0003824

(catalytic activity), GO:0004601 (peroxidase activity), GO:0043169 (cation binding), GO:0005509 (calcium ion binding), GO:0020037 (heme binding), GO:0043565 (sequence-specific DNA binding), transferase activity-transferring acyl groups (GO:0016746), transcription factor activity (GO:0003700) and GO:0016762 (xyloglucan-xyloglucosyl transferase activity) was highly expressed in the tolerant genotypes but reverse in sensitive genotypes **(Fig. 5B & 6B)** . In the CC category, plant-type vacuole membrane (GO:0000325), apoplast (GO:0048046), extracellular matrix (GO:0031012), and cell wall (GO:0005618) GO terms were up-regulated in tolerant genotypes, whereas the reverse trend was observed in sensitive genotypes**(Fig. 5C & 6C)** .

Pathway Analysis of DEGs Using MapMan

Metabolic pathways analysis of DEGs was performed through MapMan to determine the effects of drought stress in root tissues of soybean. In our data, out of all up- and down-regulated DEGs of DTP and DSP, only 121 and 105 DEGs, respectively were mapped, and out of them only 99 and 87 were shown in **Fig. 7B** , whereas only 123 and 123 DEGs were mapped out of all up- and down-regulated DEGs of DTL and DSL, respectively, and 101 and 101were presented in the **Fig. 7A** by using Gmax_AFFY_09.m02 (Thimm et al., 2004). Comprehensive overview of DEGs, such as peroxidases, TFs, secondary metabolites, PRs, cell wall, signaling, hormone signaling, redox state etc. are presented in**Fig. 7A&B** . Both MapMan and GO analysis are in well agreement, as in both cases DEGs enrichment were found similar in pathways and functional groups. However, most of the DEGs were assigned to specific functional groups, genes relevant to the terms hormone signaling, cell wall, antioxidant activity, signaling, proteolysis, TFs (MYB, ERF, WRKY, bZIP, and bHLH) were highly expressed in tolerant genotypes, and reverse trend was observed in sensitive genotypes.

Hence, consistency of GO and MapMan analyses revealed that drought stress response was highly enriched in the DEGs associated with water transport (GO:0006833), cell wall (GO:0005618), plant-type vacuole membrane (GO:0000325), antioxidant activity (GO:0016209), peroxidase activity (GO:0004601), catalytic-activity (GO:0003824), auxin transport (GO:0060918), protein-kinase activity (GO:0004672), secondary metabolite synthesis (GO:0044550), transcription-factor activity (GO:0003700), hormone-signaling (GO:0009755). Hence, these terms can play vital role in regulating drought stress tolerance at germination stage in soybean.

KEGG Pathway Analysis of DEGs Involved in Secondary Metabolism and Hormone-Mediated Signaling

Based on the GO and MapMan analysis as well as previous literature, secondary metabolites and hormone-mediated signaling plays important role in drought tolerance. Hence, DEGs related to secondary metabolites and hormone-mediated signaling were further subjected to KEGG pathway enrichment analysis. Our results showed that genes associated with biosynthesis of secondary metabolites such as shikimate acid, alkaloid, anthocyanin, lignin, flavonoids and terpenoid were mainly differentially expressed under drought stress in root tissues **(Table 2).** The number of up- and down-regulated genes associated with secondary metabolites biosynthesis such as shikimate acid, alkaloid, anthocyanin, lignin and terpenoid in both tolerant and sensitive genotypes are presented (**Fig. S1; Table S8).**

Furthermore, many DEGs related to AUX, ABA, GA, JA, ETH and Brassinosteroid (BR) were found in drought stressed soybean roots. The changes in the gene expressions of hormone-responsive genes, and their regulating TFs are shown in **Table 3.** To further understand the regulation of phytohormone in soybean transcriptome, several DEGs related to the biosynthesis of phytohormones viz., AUX, ABA, GA, JA, ETH and BR were identified under drought stress in the soybean root tissues. The changes in the gene expressions of hormone-responsive genes, and their regulating TFs are presented (**Table 3; Fig. S2)** . The number of DEGs involved in the hormonal-signaling were mostly related to AUX followed by GA, ETH and ABA **(Table S13).** This indicates important role of these hormones especially AUX in drought stress response at germination stage.

PPI Network Analysis for Differentially Expressed TFs

In the present study, 354, 348, 363 and 324 differentially expressed TFs were observed in DTL, DSL, DSP and DTP genotypes, respectively under drought stress. A total of 265 and 270 TFs were down-regulated in DSP and DSL, respectively, whereas 185 and 155 TFs were up-regulated in DTL and DTP, respectively **(Table 4).** Interestingly, the myriad of TFs, were down-regulated in sensitive genotypes (DSP & DSL) under stress **(Table 3; Table S10).** Most of the TFs were associated with MYB (myeloblastosis), bHLH (basic helix-loop-helix), ERF (Ethylene-responsive factor), NAC and WRKY families **(Table 4).**However, MYB represented the highest number of differentially expressed TFs followed by WRKY, ERF, bHLH and NAC under drought stress. To explore the interaction of drought related TFs, we constructed PPI network by using STRING. Among all the differentially expressed MYB genes, MYB127 and MYB184 showed highest and lowest expression level in tolerant and sensitive genotypes, respectively; similarly, MYB78 showed lowest expression in tolerant relative to sensitive genotypes. Hence, we checked separately the interaction of these three MYB genes with all other TFs to understand their PPI interaction with other differentially regulated TFs. Interestingly, all these three MYB genes were observed to act as hub gene interacting with the maximum number of nodes**(Fig. 8).** For example, MYB78 acting as a hub gene that interacts with NAC2, BZIP48, DREB1 and C2H2 **(Fig. 8B**). Similarly, hub gene MYB184 interacts with WRKY25, BHLH and AP2 (**Fig. 10C)** . MYB127 interacts with MYB54 and MYB48, and showed highest level of expression among all MYB in tolerant genotype suggesting its positive role in drought tolerance **(Fig. 8A).**Hence, our results propose that MYB78, MYB127 and MYB184 genes might be the key elements of stress-related signaling networks.

Candidate Gene Analysis

By taking into consideration the results of the GO, MapMan and PPI network, stress responsive DEGs having function relevant to the cell wall, plant-type vacuole membrane, water transport, antioxidant activity, auxin transport, peroxidase activity, protein kinase activity, TF activity, secondary metabolite synthesis, hormone signaling and signaling were screened (**Table 2**). A total number of 981 DEGs associated with above-mentioned functions were screened, and used for candidate gene prediction analysis. Gene ID, annotation and expression pattern of 981 genes in all four genotypes viz., DTP, DSP, DTL & DSL are presented **(Table 2; Table S7-S13)** . These all 981 DEGs were subjected to cluster analysis by using dChiP software, and were grouped into six different clusters viz., I, II, III, IV, V & VI on the basis of expression pattern of DEGs in four different genotypes **(Fig. 9).** Out of these six clusters, the DEGs in Cluster II & III showed consistent opposite expression trend between tolerant (DTP & DTL) and sensitive (DSP & DSL) genotypes. For example, the DEGs of cluster II revealed up-regulation in tolerant genotypes (DTP & DTL) and down-regulation in sensitive genotypes (DSP & DSL), whereas the DEGs of cluster III showed down-regulation in tolerant genotypes but up-regulation in sensitive genotypes. The cluster II & III include 547 DEGs out of total 981 DEGs, and these 547 genes showed consistent and opposite expression pattern in tolerant and sensitive genotypes (**Fig. 9; Table 4**), suggesting vital role of these genes in the regulation of drought tolerance in soybean.

To further predict the most possible genes among above 547 DEGs, the Chr.08 earlier reported to be rich in QTLs and "Hotspot regions" associated with drought stress tolerance at different growth stages in soybean such as germination and early seedling (V1) were screened for these DEGs (Wang *et al* ., 2012; Manavalan *et al* ., 2015; Valliyodan *et al* ., 2016). By taking this into consideration, we screened all the 547 DEGs, and 28 of them were found on Chr.08**(Table S14).** These 28 genes were further subjected for sequence polymorphism (such as SNP, InDels) detection between drought-tolerant (DTL) and drought-sensitive (DSL) lines. Out of these 28 genes, only eight genes showed non-synonymous SNPs viz., *Glyma08G162700* ,*Glyma08G091400, Glyma08G097300, Glyma08G045000, Glyma08G271600* ,*Glyma08G014200* , *Glyma.08G201700* and *Glyma08G042100* . Among these eight genes, two genes viz., *Glyma08G091400* and*Glyma08G097300* were located within "Hotspot region" on Chr.08 related with drought tolerance in soybean (Manavalan *et al* . 2015; Valliyodan *et al* . 2016)**.** The function annotation of eight plus two hub genes (*Glyma.08G029400* /MYB127 and*Glyma.10G184500* /MYB78) are presented in **Table 5.**

Hence, based on the PPI, GO, dChiP, gene annotation and sequencing analysis, these above ten genes were predicted as candidate genes regulating drought tolerance at germination stage in soybean (**Table 5**).

Therefore, further validation of these genes is required before they can be successfully employed for molecular breeding of drought tolerance in soybean.

qRT-PCR Validation

In order to validate the reliability and reproducibility of RNA-seq results, we selected above ten predicated candidate genes plus three random DEGs related to stress response for qRT-PCR validation (**Table S1**). By comparing qRT-PCR and RNAseq data, the relative expression of qRT-PCR analysis was also converted into FC by base 2 for consistency with the RNA-Seq data. These 12 genes include both up-regulated and down-regulated genes **(Table S14**). The results of qRT-PCR analysis of all these 12 genes showed consistency with the RNAseq data **(Fig. 10).** Overall, the results of qRT-PCR analysis showed the same expression trends as that of RNA-Seq, indicating the reliability of transcriptomic data used for the analysis in the present study.

Discussion

Drought stress is a leading constraint affecting grain yield and quality in soybean (Fried *et al.* , 2019). Genetic basis and mechanisms underlying drought tolerance at germination stage has not been extensively studied in soybean. Tolerance to drought stress is very complex; hence, elucidation of molecular mechanism regulating drought tolerance has been the long-term interest of soybean breeders. Identification, cloning and exploitation of stress responsive genes using molecular breeding/transgenic techniques is essential to develop drought tolerant soybean cultivars. Although, linkage and association mapping has identified genomic locus underlying drought tolerance (Semagn *et al.* , 2013; Wang and Qin, 2017), but attempts to understand drought tolerance at the transcriptome level was limited especially at germination stage. In this context, the present study used RNA-seq analysis to explore the global transcriptome of four contrasting genotypes viz., two tolerant (DTP & DTL) and sensitive (DSP & DSL) genotypes under drought stress at germination stage. Our results revealed that tolerant genotypes (DTP & DTL) possess higher RL compared to sensitive genotypes under stress treatment. This can be explained by the fact that RSA are important sensors of drought tolerance, and increased root length and deep root system greatly increases the moisture absorption and nutrient extraction for plant survival under water deficit stress (Wasaya *et al.* , 2018).

Both Venn diagram and Circos analysis suggests that sensitive genotypes are comparatively more vulnerable to drought stress at the transcriptomic level (**Fig 2A, B & C**). Similar findings are previously reported in several other crop plants under water-deficit and other stresses (Muthusamy *et al.* , 2016; Fracasso *et al.* , 2016). These results can be explained by the fact that compared to tolerant genotypes, sensitive genotypes reveal dramatic changes in morpho-physiological and biochemical parameters while mitigating negative impact of drought stress (Yang *et al.* , 2017). The higher correlation of DTP with DTL compared to DSP and DSL provides explanation about increased drought tolerance in DTL, because genes/alleles for drought tolerance in DTL are derived from wild-parent/DTP, while same genes/alleles for DSL comes from DSP (**Fig. 3**).

Both GO and MapMan analyses indicated that drought stress response was highly enriched in the DEGs associated with water transport, cell wall, plant-type vacuole membrane, antioxidant activity, catalytic activity, auxin transport, peroxidase activity, protein kinase activity, TF activity, secondary metabolite synthesis, hormone signaling and signaling (**Table 4**). Under water deficit conditions, plant cells must maintain functional integrity and rapidly remodeled to keep cell wall flexible under abiotic stress (Houston *et al.* , 2016). In this context, cell-wall remodeling enzymes such as pectin esterases (PME), expansins, xyloglucan endo-transglucosylase/hydrolase (XTH) as well as glycine-rich cell wall structural protein (GRP) are involved in maintaining cell wall rigidity, and permits to cope with water deficit conditions (Tenhaken, 2015). However, most of these cell wall remodeling enzymes were highly expressed in tolerant genotypes but down-regulated in sensitive ones, indicating their vital role in cell wall remodeling under drought stress (**Table 4**). For instance, Arabidopsis lines overexpressing *PMEI1* (pectin methylesterases) exhibited improved germination rate and seedling root growth under water deficit condition (An et al. 2008). Cho *et al.* , (2006) described that *CaXTH3*overexpressing-plants in *Arabidopsis* showed an increase in drought tolerance, and *XTHs* have

been reported to be differentially regulated in maize under drought stress (Zhu *et al.* , 2007). Furthermore, overexpression of expansin gene (*TaEXPB23*) improves drought tolerance in tobacco. Similarly,*AtEXP2* is involved in seed germination and drought stress response in *Arabidopsis* (Yan *et al.* , 2014). Yang *et al.* , (2014) also reported that overexpression of *AtGRP2*and *AtGRP7* significantly influences drought tolerance in transgenic rice, and revealed that rice plants overexpressing*GRPs* were more tolerant to water deficit relative to wild-type plants. Xuan *et al.* , (2010) also demonstrated role of*NtGRP-1a* transcripts under drought stress in tobacco. Hence, up-regulation of above cell wall related genes might play essential role in seed germination and drought tolerance.

Under stress conditions different secondary metabolites i.e., lipids, amino acid and carbohydrate are accumulated in higher plants (Akula and Ravishankar, 2011). Shikimate pathway not only act as connection between central and secondary metabolism but also serve as precursor for most of the other secondary metabolites (**Fig. S1** ; Maeda and Dudareva, 2012). Biosynthesis of tyrosine and phenylalanine through the shikimate pathway leading to the synthesis of wide range of secondary metabolites (Less and Galili, 2008; Gill and Tuteja, 2010). Isoquinoline alkaloids are derived from tyrosine, while indole alkaloids are produced by metabolic engineering of the tryptophan preventing plants from oxidative stress (**Figure. S1** ; Sato and Kumagai, 2013). Phenylalanine acts as a precursor for diverse secondary metabolites and phenylalanine ammonia-lyase (PAL) takes part in phenylpropanoid biosynthesis; an essential step towards biosynthesis of anthocyanin's, flavonoids, stilbenes, lignins and other compounds (Deng and Lu, 2017). It is reported that lignin and anthocyanin's biosynthesis related genes were highly expressed under water stress in rice roots (Yang *et al.* , 2006). In our findings, many unique genes related to biosynthesis of lignin, alkaloids, flavonoids and anthocyanin's such as PAL, STR, and laccase were highly expressed in tolerant genotypes, but mostly down-regulated in sensitive genotypes under drought (**Table 4**), proposing their diverse role towards drought tolerance. Among all the DEGs involved in the secondary metabolite biosynthesis, highest number of genes associated with lignin biosynthesis especially the laccase (LAC) and cinnamoyl-CoA reductase (CCR) genes indicating essential role of lignin in drought tolerance. Lignin reduces cell wall water penetration and transpiration in plants, and also maintain cell osmotic balance and membrane integrity under drought stress (Liu *et al* ., 2018). Moreover, laccase genes participate in the oxidation of flavonoids, and plays fundamental role in plant defense responses (Turlapati *et al.* , 2011). Flavonoids serve as antioxidant, and provide shield to plants against abiotic stresses (Pourcel *et al.* , 2007). For example, *AtLAC2* in*Arabidopsis* is involved in drought stress tolerance (Cai *et al.* , 2006). Liu *et al.* , (2017) reported accumulation of terpenoid indole alkaloids (TIAs) in the *C. roseus* under drought stress. Hence, considerable upregulation of DEGs related to secondary metabolite biosynthesis in drought-tolerant genotypes, suggesting their vital role in drought tolerance.

Drought stress enhances reactive oxygen species (ROS) accumulation resulting in cell wall degradation and membrane damage.(Helena and Carvalho, 2008). Studies have demonstrated that ROS accumulation impaired seed germination in different crops including soybean (Ishibashi *et al.* , 2013). In this regard, plant utilize antioxidant defense system for scavenging ROS to avoid oxidative stress, and allows proper seed germination (Xie *et al.* , 2019). In our data, many redox related DEGs including GST, POD, SOD, glutaredoxin and thioredoxin were highly expressed in the drought-tolerant genotypes (DTP and DTL), whereas the same genes were down-regulated in drought-sensitive genotypes viz., DSP and DSL (**Table 4**). Many studies have revealed induced expression of GST under various abiotic stresses, including drought (Kumar and Trivedi, 2018). For instance, Bhardwaj and Yadav, (2012) reported an increase in expression of GST, POD, CAT, and SOD in horsegram subjected to drought stress, indicating their important role in drought stress tolerance. In addition, over-expression of *AgAPX1* (ascorbate peroxidase) in Arabidopsis led to increase in ascorbate content and drought tolerance (Liu *et al.* , 2019). A glutaredoxin gene viz., *SlGRX1* exhibited tolerance to oxidative and drought stresses in tomato (Guo *et al.* , 2010). Similarly, NADPH-dependent thioredoxin reductase A (*NTRA*) mutant exhibited tolerance to drought and oxidative stress (Cha *et al* . 2014).

In soybean and other crop species, many transcriptome studies has revealed involvement of several TFs including, MYB, ERF/DREB, bHLH, NAC and WRKY in seed-germination process under stress conditions (Lee*et al.* , 2015; Raineri *et al.* , 2016; Baillo *et al.* , 2019). Members of these TF families may have

either positive and negative regulatory role in the drought stress tolerance. The MYB are most abundant among all TFs expressed in present study, and many of these TFs are differentially regulated in both tolerant and sensitive genotypes (**Table 3**), suggesting their essential roles in germination in response to drought stress. For example, *OsMYB6*overexpressing plants in rice enhanced tolerance to drought stress (Tang*et al.* , 2019). Similarly, Zhao *et al.* , (2018) reported the involvement of *TaMYB* towards drought tolerance in *Arabidopsis* . Overexpression of the *ZmMYB3R* and*ScMYBAS1* has been reported to enhance drought stress tolerance in maize (Wu *et al.* , 2019) and rice (Peixoto-junior *et al.* , 2018), respectively. After MYB, the bHLH followed by ERF, NAC and WRKY represent the highest number of differentially expressed TFs under drought stress. For example, bHLH family member *ZmPTF1* and*VvbHLH1* regulates drought tolerance in maize (Li *et al.* , 2019) and *Arabidopsis thaliana* (Wang, *et al.* , 2016), respectively by promoting root development, ABA synthesis and accumulation of flavonoids. The AP2/ERF TF were revealed to modulate brassinosteroid-regulated plant development and drought responses in*Arabidopsis* (Xie *et al.* , 2019). Moreover, NAC genes such as *JUNGBRUNNEN1* in tomato (Thirumalaikumar *et al.* , 2018) and *OoNAC72 Arabidopsis* (Guan *et al.* , 2019), positively regulates drought tolerance, whereas, *SbNAC052* , *SbNAC073* , and *SbNAC116* serve as negative regulator in drought stress tolerance in sorghum (Sanjari *et al.* , 2019). Many TFs from WRKY family exhibited function in drought tolerance in various crops, for example overexpression of *GmWRKY12* in soybean (Shi *et al.2018*), *TaWRKY2* and *AtWRKY30* in wheat (Gao *et al.* , 2018; El-Esawi et al. 2019), and *ZmWRKY40* in Arabidopsis (Wang *et al.2018*) has led to enhanced drought tolerance in these crop plants. Hence, these TFs can be an important target for breeding drought tolerance in soybean.

In the present study, many genes related to plant-specific Ca^{2+} signaling such as calcium-binding proteins, calcium ATPases, calmodulin-like proteins (CMLs), calmodulin-binding protein, calmodulin-binding receptor, Ca^{2+}-dependent protein kinases (CPKs), and Annexin were mostly up-regulated in tolerant genotypes as compared to sensitive genotypes. (**Table 4**). For example, Campo *et al.* , (2014) reported that *OsCPK4*overexpression in rice showed an increase in drought tolerance by reducing lipid peroxidation and electrolyte leakage. Plasma membrane Ca^{2+}-ATPase directly regulates drought stress tolerance by activating ABA signaling pathway (Shao *et al.* , 2008; Cerana *et al.* , 2006) and the increased ABA accumulation leads to stomatal closure and expression of many stress-related genes. For example, Cerana *et al* ., (2006) reported overexpression of Ca^{2+}-ATPase viz.,*ACA8* and *ACA9* stimulated ABA accumulation. Hence, it has been proposed that Ca^{2+}-ATPases might play important role in drought stress response through ABA signaling, as the latter is well-established mediator of drought stress adaptation in plants (Qudeimat *et al.* , 2008). Calmodulin-like Proteins such as*CML20* in *Arabidopsis* (Wu *et al.* , 2017),*ShCML44* from *Solanum habrochaites* (Munir *et al.* , 2016) were revealed to induce drought tolerance by regulating ABA signaling in guard cells. The annexin gene *AnnSp2* exhibited drought tolerance in overexpressed transgenic tomato plants through ROS-scavenging and modulation of ABA synthesis (Ijaz *et al.* , 2017). Wei *et al.,* (2014) reported that *OsCPK9,*calcium-dependent protein kinase in rice is involved in tolerance to drought stress. Hence, the above findings suggest close relationship between ABA and Ca-signaling in drought stress response. Evidence indicate that Ca^{2+}/CaM is involved in ABA-induced drought signaling under PEG stress, and ABA-synthesis was associated with cytoplasmic Ca^{2+} concentrations (Li *et al* . 2002). As reported previously, ABA activates cytosolic Ca^{2+} in guard cells by maintaining turgor within guard cells that leads to stomatal closer and prevention of transpiration water loss, and ultimately induces drought tolerance (Song *et al.* , 2008). Furthermore, receptor-like protein kinase (RLKs) and mitogen-activated protein kinases (MAPKs) are key components for signaling pathways in plant and have diverse function in seed germination by regulating stress-responsive gene (Baek *et al.* , 2019; Jagodzik *et al.* , 2018). For instance, seed germination was significantly enhanced by expression of *SpMAPK3* and*SlMAPK3* in response to abiotic stresses (Muhammad *et al.* , 2019; Li *et al.* , 2014). The *MAPK kinase10.2* promotes drought tolerance by activating different *MAPKs* in rice (Ma*et al.* , 2017). Cysteine-rich RLKs such as CRK45 in*Arabidopsis* (Zhang *et al.* , 2013) and *TaCRK41* in wheat (Chen *et al.* , 2017) involved in ABA signaling and positively regulates seeds germination under drought and oxidative stress. Similarly, in Arabidopsis, proline-rich RLKs, *PnLRRRLK27*exhibited tolerance towards abiotic stress during seed germination (Wang*et al.* , 2017). Consistent with these results, we also observed up-regulation of MAPKs and RLKs in tolerant genotypes, whereas,

reverse trend was seen in sensitive genotypes. Therefore, the above results propose that MAPKs, RLKs and Ca^{2+} signaling together with ABA might play key role to regulate seed germination under water deficit conditions.

Numerous transgenic studies have demonstrated that overexpression of Aquaporins (AQPs) viz., TIPs and PIPs enhanced drought tolerance probably by endorsing stomatal closure and regulating the plant hydraulics (Zargar *et al.* , 2017). For example, *VfPIP1* overexpressing plants in *Arabidopsis thaliana* exhibits drought tolerance through promoting stomatal closure (Cui *et al.* , 2008). Pou *et al.* , (2013) reported the putative role of *PIP* and *TIP* genes in leaf hydraulic and stomatal conductance in grapevine under drought stress. Although role of AQPs in dehydration tolerance has been extensively studied in plants but still conflict of interest remained among the researchers regarding their up- or down-regulation during water deficit condition. Under drought stress, considerable variation was observed in the expression of *PIPs* at the transcript level; for example, significant upregulation of some *PIPs* was observed while expression of others remained either unchanged or downregulated (Lian *et al.* , 2006; Guo *et al.* , 2006). Upregulation of some AQPs might assist in maintaining the normal physiological processes in plant and resist the stress, while other may help to adapt or tolerate the stress condition by reducing their own activity and expression (Zargar *et al* . 2017). Hence, in the present study some unique AQP genes are differentially regulated in tolerant and sensitive genotypes proposing their innate role in drought tolerance (**Table 4**).

Phytohormone signaling is complex and plays important regulatory role in drought responsive pathways of soybean (Pandey *et al* , 2017). In recent study, genes associated with phytohormone biosynthesis and signaling viz., AUX, ETH, ABA and GA represent most number of DEGs in both tolerant and sensitive genotypes under drought stress (**Table 2; Table 4**). For example, auxin/indole-3-acetic acid (Aux/IAA), small auxin-up RNA (SAUR), indole-3-acetic acid-amido synthetase/Gretchen Hagen (GH3) and auxin efflux carrier component (PIN) showed dynamic changes in the roots of tolerant and sensitive genotypes under drought stress; however, these genes are highly expressed in tolerant genotypes, whereas, reverse trend was observed in sensitive genotypes, indicating complex role of auxin signaling. Zhang *et al.* (2012) revealed that auxin efflux carrier component such as *OsPIN3t* in rice involved in auxin transport and response to water stress. The auxin-sensitive Aux/IAA proteins has been revealed to mediate drought tolerance through regulating glucosinolate levels in Arabidopsis (Salehin *et al.* , 2019). Under drought stress, overexpression of auxin-responsive protein, *TaSAUR75* exhibited an increase in root length and survival rate in *Arabidopsis*(Guo *et al.* , 2018). Similarly, *GH3* genes of Chickpea (*CaGH3-1 & CaGH3-7*) and *Medicago truncatula*(*MtGH3-7* , *MtGH3-8* and *MtGH3-9*) were highly induced under drought stresses (Singh *et al.* , 2015). In our findings, many *GH3* genes were highly induced in tolerant genotypes, respectively, suggesting their important role in stress adaptation. Ethylene biosynthesis mainly involved two enzymes viz., aminocyclopropane-1-carboxylic acid (ACC) oxidase and ACC-synthase (ACS) in various crops (Yoon and Kieber, 2013; Van Der Straeten *et al* ., 2001), and this hormone regulates seed germination and seedling growth under abiotic stresses (Huang *et al.* , 2019). Similar with these findings, our study also exhibited higher expression levels of ACS and ACO in tolerant genotypes. In addition, it is also reported that hormonal balance between ABA and GAs is necessary to regulate seed germination (Miransari and Smith, 2014). In present study, DEGs associated with gibberellin 2-beta-dioxygenase, Gibberellin 3-beta-dioxygenase, Gibberellin 20 oxidase, and DELLA protein were highly expressed in tolerant genotypes as compared to sensitive genotypes **(Table 1).** Habib *et al.* , (2019) demonstrated that *SlGRAS7* TF improved drought tolerance by enhancing gibberellin/auxin signaling. Consistent with a recent report, we also identified many enzymes related to ABA biosynthesis such as 9-cis-epoxycarotenoid dioxygenase (NCED), abscisic acid 8'-hydroxylase and carotenoid cleavage dioxygenases (CCD). These enzymes were differentially regulated in tolerant and sensitive genotypes resulting in better hormonal balance under drought stress (Shu *et al.* , 2018). In plants, ABA plays diverse role in response to drought as well as in various developmental processes such as seed germination. Drought stress enhances plant ABA accumulation resulting in ABA-receptor complex (PYR/PYL/RCAR–PP2C–ABA) that triggers SnRK2 protein kinases, and this kinase facilitate stomatal closure and reducing transpiration water loss, and maintaining water balance within plant under water-deficit condition (Fujita *et al.* , 2013). Brassinosteroids (BRs) are plant steroid hormones, which regulate the expression of stress

related genes. Brassinosteroid Insensitive 1 (BRI1) was highly expressed in tolerant genotype, and previous study has revealed that *BRI1* plays essential role in plant growth, development and drought tolerance. For example, Feng *et al.* , (2015) using RNAi suggested that mutation of *bdBRI1* in *Brachypodium distachyon* resulting in dwarf phenotype with enhanced tolerance towards drought stress.

Based on the findings of GO enrichment, PPI network and dChiP analysis as well as screening of genes within previously reported genomic/hotspot regions associated with drought tolerance in soybean on Chr.08, a total of ten genes were considered as the most possible candidate genes regulating drought tolerance at germination stage in soybean (**Table 5**). Function annotation revealed that*Glyma08G162700* function as peroxidase 3 (POD 3),*Glyma08G091400* as 'glutamate decarboxylase 1-like',*Glyma08G097300* as *peroxisomal (S)-2-hydroxy-acid oxidase GLO1-like* , *Glyma08G045000* as Ca-transporting ATPase,*Glyma08G271600* function as glycine-rich cell wall structural protein, *Glyma08G014200* function as tubulin beta-1 chain-like,*Glyma.08G201700* function is uncharacterized; the remaining three predicted genes viz., *Glyma08G042100* , *Glyma.08G029400* and*Glyma.10G184500* function as *MYB184, MYB127* and*MYB78,* respectively **(Table 5)** . The POD 3 enzymes were reported to participate in plant development, stress responses and hormone signaling (Wu *et al.* , 2019); for example, *AtPrx3*participates in positive regulation drought stresses response in Arabidopsis (Llorente *et al.* , 2002). Yong *et al* . (2017) reported that inhibition of glutamate decarboxylase activity result in the increase of endogenous glutamate (Glu), and that in turn enhanced drought tolerance in white clover. Down-regulation of *glycolate oxidase* pathway reduces peroxisomal H_2O_2 production in the green tissues of plants under drought stress, and provides tolerance to oxidative stress (Zhou *et al* ., 2007; Noctor *et al* ., 2014). The Ca^{2+}ATPase such as *OsACA6* confers drought stress tolerance with reduced accumulation of ROS and enhanced the expression of stress-responsive genes in tobacco (Huda *et al.* , 2013). Involvement of Ca^{2+}-ATPase in drought tolerance has been revealed through transcript profiling of a sweet potato (Yang*et al.* , 2018). Moreover, in *B. napus* , two cell-wall related proteins viz., glycine-rich and fasciclin-like arabinogalactan were induced under drought stress (Koh *et al.* , 2015). Pan*et al* . (2018) reported highly induced expression of tubulin beta-1 chain-like in foxtail millet under drought stress (Pan *et al.* , 2018). The tubulin beta-1 (R83) chains are major constituents of microtubules and their accumulation also peaked at 48 h after the onset of drought stress (Bian *et al.* , 2017). In addition, many of the earlier studies has explained the important role of MYB gene family in the drought stress tolerance in crop plants (Tang *et al* ., 2019; Zhao *et al* ., 2018; Wu *et al* ., 2019; Zhang *et al* ., 2019; Peixoto-Junior *et al* ., 2018). Hence, the above ten genes that includes three hub genes were predicted as the most possible candidate genes regulating drought tolerance at germination stage in soybean. Therefore, these genes required further validation to prove their actual role and use in soybean improvement.

Based on the above findings, a hypothetical model was proposed as shown in **Fig. 11** . This comparative model explains, how aquaporin's, cell-wall related enzymes, secondary metabolites, antioxidants, kinases, MAPK signaling and TF activities functions in drought response between drought tolerant and sensitive genotypes.

Conclusion

The current study presents detailed information on drought tolerance at germination stage in four contrasting soybean genotypes viz., drought-tolerant (DTP & DTL) and drought-sensitive (DSP& DSL), and revealed the molecular mechanisms of drought tolerance. Our transcriptomic data revealed that tolerant genotype (DTP & DTL) is less affected compared to the sensitive genotypes by drought stress. Moreover, out of 981 DEGs screened for candidate gene prediction analysis, 547 exhibited significant opposite gene expression pattern between tolerant and sensitive genotypes, and 28 of them were located on the Chr.08 reported to possess most number of drought tolerant QTLs/"QTL Hotspots". Eight of these 28 genes showed non-synonymous SNPs, and two were located within "Hotspot region" associated with drought tolerance on Chr.08. Hence, based on the gene annotation, PPI and sequencing analysis, ten genes including above eight genes plus two hub genes were predicated as most possible candidate genes regulating drought tolerance in soybean. Hence, these ten genes required functional validation before they can be used for breeding drought tolerance in soybean. Overall, our study provides valuable information to understand drought tolerance mechanism, and

will greatly assists in cloning drought-tolerant genes for breeding improved soybean cultivars.

Acknowledgements

This research was funded by the National Key Research and Development Program (2016YFD0100201, 2018YFD0201006), the National Natural Science Foundation of China (Grant Nos. 31871646 and 31771821), the MOE Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT_17R55), the Fundamental Research Funds for the Central Universities (KYT201801), and the Jiangsu Collaborative Innovation Center for Modern Crop Production (JCIC-MCP) Program.

Author Contributions

TJZ, JAB conceived and designed the experiments. MA performed the experiments. MA, MMR, MSH, RMA, ZA analyzed the data. MA drafted the manuscript. TJZ, JAB revised the paper.

Competing interest

The authors declare they have no competing interests.

Ethics and consent to participate

This study did not involve humans, human data or animals; no ethics approval or consent is required to publish the results.

Availability of data and materials

All generated data and material can be accessed at any time.

References:

Ahmad, S., Ahmad, R., Ashraf, M.Y., Ashraf, M. and Waraich, E.A. (2009) Sunflower (*Helianthus annuus* L.) response to drought stress at germination and seedling growth stages. Pakistan Journal of Botany. **41,** 647-654.

An, S.H., Sohn, K.H., Choi, H.W., Hwang, I.S., Lee, S.C., and Hwang, B.K. (2008). Pepper pectin methylesterase inhibitor protein*CaPMEI1* is required for antifungal activity, basal disease resistance and abiotic stress tolerance. *Planta* , 228: 61-78.

Baek, D., Kim, M.C., Kumar, D., Park, B., Cheong, M.S., Choi, W.,*et al* . (2019) *AtPR5K2, a PR5-* Like Receptor Kinase, Modulates Plant Responses to Drought Stress by Phosphorylating Protein Phosphatase 2Cs. Frontiers in Plant Science, **10** , 1146.

Baillo, E.H., Kimotho, R.N., Zhang, Z. and Xu, P. (2019) Transcription factors associated with abiotic and biotic stress tolerance and their potential for crops improvement. *Genes* , **10** , 771.

Bailey-Serres, J., Lee, S.C., and Brinton, E. (2012). Waterproofing crops: effective flooding survival strategies. *Plant Physioloy* , **160** , 1698-1709.

Bewley, J.D., Bradford, K.J., Hilhorst, H.W.M. and Nonogaki, H. (2013). Seeds: Physiology of Development, Germination and Dormancy. 3rd edn. New York: Springer. doi: 10.1007/978-1-4614-4693-4.

Bhardwaj, J. and Yadav, S.K. (2012) Comparative study on biochemical parameters and antioxidant enzymes in a drought tolerant and a sensitive variety of horsegram (*Macrotyloma uniflorum*) under drought stress. *American Journal of Plant Physiology* ,7, 17-29.

Bian, Y., Deng, X., Yan, X., Zhou, J., Yuan, L. and Yan, Y. (2017) Integrated proteomic analysis of Brachypodium distachyon roots and leaves reveals a synergistic network in the response to drought stress and recovery. *Scientific Reports* , **7** , 46183.

Bolger, A.M., Lohse, M. and Usadel, B. (2014) Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* , **30** , 2114–2120.

Boureima, S., Eyletters, M., Diouf, M., Diop, T.A. and Van Damme, P. (2011). Sensitivity of seed germination and seedling radicle growth to drought stress in sesame (*Sesamum indicum* L.). *Research Journal of Environmental Science* , **5** , 557–564.

Cai, X., Davis, E. J., Ballif, J., Liang, M., Bushman, E., Haroldsen, V., *et al* . (2006). Mutant identification and characterization of the laccase gene family in Arabidopsis. *Journal of Experimental Bot* any, **57** , 2563–2569

Campo, S., Baldrich, P., Messeguer, J., Lalanne, E. and Coca, M. (2014) Overexpression of a Calcium-Dependent Protein Kinase Confers Salt and Drought Tolerance in Rice by Preventing membrane lipid peroxidation. *Plant Physiol* ogy, **165** , 688-704.

Cerana, M., Bonza, M.C., Harris, R., Sanders, D. and De Michelis, M.I. (2006) Abscisic Acid Stimulates the Expression of Two Isoforms of Plasma Membrane Ca^{2+}-ATPase in *Arabidopsis thaliana*Seedlings. *Plant Biology,* **8** , 572–578.

Chaudhary, J., Patil, G.B., Sonah, H., Deshmukh, R.K., Vuong, T.D., Valliyodan, B. and Nguyen, H.T. (2015) Expanding omics resources for improvement of soybean seed composition traits. *Frontiers in Plant Science* , **6** , 1021.

Cha, J.Y., Kim, J.Y., Jung, I. J., Kim, M.R., Melencion, A., Alam, S. S., et al (2014). NADPH-dependent thioredoxin reductase A (*NTRA*) confers elevated tolerance to oxidative stress and drought. *Plant Physiology and Biochem* istry, **80** , 184–191.

Chen, D., Wu, J., Zhao, M., Ma, X., Zhang, W., Xia, G. and Wang, M. (2017) A novel wheat cysteine-rich receptor-like kinase gene*CRK41* is involved in the regulation of seed germination under osmotic stress in *Arabidopsis thaliana* . *Journal of Plant Biology* **60** , 571–581.

Chen, W., Yao, Q., Patil, G.B., Agarwal, G., Deshmukh, R.K., Lin, L.,*et al* . (2016) Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. *Frontiers in Plant Science,***7** , 1044.

Cho, S.K., Kim, J.E., Park, J.A., Eom, T.J. and Kim, W.T. (2006) Constitutive expression of abiotic stress-inducible hot pepper*CaXTH3* , which encodes a xyloglucan endotransglucosylase/hydrolase homolog, improves drought and salt tolerance in transgenic Arabidopsis plants. *FEBS Letters.* **580** , 3136-3144.

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X. and Ruden, D.M. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly, 6, 80–92

Collard, B.C., Jahufer, M.Z.Z., Brouwer, J.B. and Pang, E.C.K. (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica* , **142** , 169-196.

Cui, X.H., Hao, F.S., Chen, H., Chen, J. and Wang, X.C. (2008) Expression of the *Vicia faba VfPIP1* gene in*Arabidopsis thaliana* plants improves their drought resistance.*Journal of Plant Research* , **121** , 207–214.

Ahmad, I.A., Din, K.U., ahmad Dar, Z., Sofi, P.A. and Lone, A.A. (2018). Effect of Drought On The Germination of Maize Using PEG (Polyethylene Glycol) As A Substitute For Drought Screening. *bioRxiv* , p.362160.

Degenkolbe, T., Do, P.T., Zuther, E., Repsilber, D., Walther, D., Hincha, D.K. and Kohl, K.I. (2009) Expression profiling of rice cultivars differing in their tolerance to long-term drought stress. *Plant Molecular Biology* , **69** , 133– 153.

Deng, Y. and Lu, S. (2017) Biosynthesis and Regulation of Phenylpropanoids in Plants. *Critical Review in Plant Science.***36** , 257–290.

Delachiave, M.E.A. and Pinho, S.Z.D. (2003) Germination of Senna Occidentalis Link : Seed at Different Osmotic Potential Levels.*Brazilian Archives of Biology and Technology* . **46** , 163–166.

El-Esawi, A. M., Al-Ghamdi, A.A., Ali, M. H. and Ahmad, M. (2019) Overexpression of *AtWRKY30* Transcription Factor Enhances Heat and Drought Stress Tolerance in Wheat (*Triticum aestivum* L.).*Genes* . 10, 163.

Feng, Y., Yin, Y. and Fei, S. (2015) Down-regulation of *BdBRI1* , a putative brassinosteroid receptor gene produces a dwarf phenotype with enhanced drought tolerance in *Brachypodium distachyon* .*Plant Science.* , **234** , 163–173.

Fracasso, A., Trindade, L.M. and Amaducci, S. (2016) Drought stress tolerance strategies revealed by RNA-Seq in two sorghum genotypes with contrasting WUE. *BMC Plant Biology.* , **16** , 115.

Fried, H.G., Narayanan, S. and Fallen, B. (2019) Evaluation of soybean [*Glycine max* (L.) Merr.] genotypes for yield, water use efficiency, and root traits. *PLoS One* , **14** , e0212700.

Fujita, Y., Yoshida, T. and Yamaguchi-Shinozaki, K. (2013) Pivotal role of the *AREB/ABF-SnRK2* pathway in *ABRE* -mediated transcription in response to osmotic stress in plants. *Physiologia Plantarum,* **147** , 15–27.

Gill, S.S. and Tuteja, N. (2010) Plant Physiology and Biochemistry Reactive oxygen species and antioxidant machinery in abiotic stress tolerance in crop plants. *Plant Physiology and Biochemistry,***48** , 909–930.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., *et al* . (2012) Phytozome: A comparative platform for green plant genomics. *Nucleic Acids Research,* **40** , D1178–D1186.

Guan, H., Liu, X., Niu, F., Zhao, Q., Fan, N., Cao, D., *et al* . (2019) *OoNAC72 , a NAC-Type* Oxytropis ochrocephala Transcription Factor , Conferring Enhanced Drought and Salt Stress Tolerance in Arabidopsis. *Frontiers in Plant Science,* **10** , 890.

Guo, L., Zi, Y.W., Lin, H., Wei, E.C., Chen, J., Liu, M., *et al* . (2006) Expression and functional analysis of the rice plasma-membrane intrinsic protein gene family. *Cell Research,* **16** , 277–286.

Guo, Y., Huang, C., Xie, Y., Song, F. and Zhou, X. (2010) A tomato glutaredoxin gene *SlGRX1* regulates plant responses to oxidative, drought and salt stresses. *Planta* , **232** , 1499–1509.

Guo, Y., Jiang, Q., Hu, Z., Sun, X., Fan, S. and Zhang, H. (2018) Function of the auxin-responsive gene *TaSAUR75* under salt and drought stress. *Crop Journal,* **6** , 181–190.

Habib, S., Waseem, M., Li, N., Yang, L. and Li, Z. (2019) Overexpression of *SlGRAS7* affects multiple behaviors leading to confer abiotic stresses tolerance and impacts gibberellin and auxin signaling in tomato. *International Journal of Genomics* , 4051981.

Helena, M. and Carvalho, C. De (2008) *Drought stress and reactive oxygen species* . **3** , 156–165.

Houston, K., Tucker, M.R., Chowdhury, J., Shirley, N. and Little, A. (2016) The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. *Frontiers in Plant Science,* **7** ,984.

Huang, G. T., Ma, S. L., Bai, L. P., Zhang, L., Ma, H., Jia, P., et al. (2012) Signal transduction during cold, salt, and drought stresses in plants. *Molecular Biology Reports,* **39** , 969–987.

Huang, Y.C., Yeh, T.H. and Yang, C.Y. (2019) Ethylene signaling involves in seeds germination upon submergence and antioxidant response elicited confers submergence tolerance to rice seedlings. *Rice* ,**12** , 23.

Huda, K.M., Banu, M.S.A., Garg, B., Tula, S., Tuteja, R. and Tuteja, N. (2013) *OsACA6* , a P-type IIB Ca^{2+}-ATPase promotes salinity and drought stress tolerance in tobacco by ROS scavenging and enhancing the expression of stress-responsive genes.*Plant Journal,* **76** , 997–1015.

Ijaz, R., Ejaz, J., Gao, S., Liu, T., Imtiaz, M., Ye, Z. and Wang, T. (2017) Overexpression of annexin gene *AnnSp2* , enhances drought and salt tolerance through modulation of ABA synthesis and scavenging ROS in tomato. *Scientific Reports* , 7, 12087.

Ishibashi, Y., Koda, Y., Zheng, S., Yuasa, T. and Iwaya-inoue, M. (2013) Regulation of soybean seed germination through ethylene production in response to reactive oxygen species. *Annals of Botany,***111** , 95–102.

Jagodzik, P., Tajdel-Zielinska, M., Ciesla, A., Marczak, M. and Ludwikow, A. (2018) Mitogen-activated protein kinase cascades in plant hormone signaling. *Frontiers in Plant Science,* **9** , 1387.

Jiménez, S., Dridi, J., Gutiérrez, D., Moret, D., Irigoyen, J.J., Moreno, M.A. and Gogorcena, Y. (2013). Physiological, biochemical and molecular responses in four *Prunus* rootstocks submitted to drought stress. *Tree Physiology,* **33** , 1061–1075.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M.,*et al* . (2008) KEGG for linking genomes to life and the environment. *Nucleic Acids Research,* **36** , 480–484.

Kim, D., Langmead, B. and Salzberg, S.L. (2015) HISAT: a fast spliced aligner with low memory requirements. *Nature Methods,***12** , 357–360.

Kobraee, S., Shamsi, K. and Rasekhi, B. (2011) Soybean production under water deficit conditions. *Annals of Biological Research* ,**2** , 423–434.

Koh, J., Chen, G., Yoo, M.J., Zhu, N., Dufresne, D., Erickson, J.E., Shao, H. and Chen, S. (2015) Comparative proteomic analysis of Brassica napus in response to drought stress. *Journal of Proteome Research,* **14** , 3068–3081.

Kumar, S., and Trivedi, P.K. (2018). Glutathione S-transferases: role in combating abiotic stresses including arsenic detoxification in plants. *Frontiers in Plant Science* , **9** .

Lan Thi Hoang, X., Du Nhi, N.H., Binh Anh Thu, N., Phuong Thao, N. and Phan Tran, L.S. (2017) Transcription factors and their roles in signal transduction in plants under abiotic stresses. *Current Genomics* ,**18** , 483-497.

Laxa, M., Liebthal, M., Telman, W., Chibani, K. and Dietz, K.J. (2019) The role of the plant antioxidant system in drought tolerance.*Antioxidants* , **8** ,94.

Le, D.T., Nishiyama, R., Watanabe, Y., Tanaka, M., Seki, M., Yamaguchi-Shinozaki, K., 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.

Lee, K., Lee, H.G., Yoon, S., Kim, H.U. and Seo, P.J. (2015) The arabidopsis *MYB96* transcription factor is a positive regulator of*ABSCISIC ACID-INSENSITIVE4* in the control of seed germination.*Plant Physiology,* **168** , 677–689.

Lenka, S.K., Katiyar, A., Chinnusamy, V. and Bansal, K.C. (2011) Comparative analysis of drought-responsive transcriptome in Indica rice genotypes with contrasting drought tolerance. *Plant Biotechnology Journal,* **9** , 315–327.

Less, H. and Galili, G. (2008) Principal Transcriptional Programs Regulating Plant Amino Acid Metabolism in Response to abiotic stresses.*Plant Physiology,* **147** , 316–330.

Li, C., Chang, P.P., Ghebremariam, K.M., Qin, L. and Liang, Y. (2014) Overexpression of tomato *SpMPK3* gene in Arabidopsis enhances the osmotic tolerance. *Biochemical and Biophysical Research Communication,* **443** , 357–362.

Li, H. and Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* **25** , 1754–1760.

Li, J.M., Cui, S.P., Guan, J.F, Li, M.J. and Guo, X.L. (2002) The Relationship between ABA and Ca^{2+}/CaM in Winter Wheat Seedlings under PEG Stress. *Acta Agronomica Sinica.* **28** , 537–540.

Li, Z., Liu, C., Zhang, Y., Wang, B., Ran, Q. and Zhang, J. (2019) The*bHLH* family member *ZmPTF1* regulates drought tolerance in maize by promoting root development and abscisic acid synthesis.*Journal of Experimental Botany* **, 70** , 5471–5486.

Lian, H.L., Yu, X., Lane, D., Sun, W.N., Tang, Z.C. and Su, W.A. (2006) Upland rice and lowland rice exhibited different PIP expression under water deficit and ABA treatment. *Cell Research* , **16** , 651–660.

Lin, M., Wei, L.J., Sellers, W.R., Lieberfarb, M., Wong, W.H., and Li, C. (2004). dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. Bioinformatics **20** , 1233-1240.

Liu, J.X., Feng, K., Duan, A.Q., Li, H., Yang, Q.Q., Xu, Z.S. and Xiong, A.S. (2019) Isolation, purification and characterization of an ascorbate peroxidase from celery and overexpression of the *AgAPX1* gene enhanced ascorbate content and drought tolerance in Arabidopsis.*BMC Plant Biology* . 19 , 1-13.

Liu, Y., Meng, Q., Duan, X., Zhang, Z. and Li, D. (2017) Effects of PEG-induced drought stress on regulation of indole alkaloid biosynthesis in *Catharanthus roseus* Effects of PEG-induced drought stress on regulation of indole alkaloid biosynthesis. *Journal of Plant Interactions.* **12** , 87-91.

Liu, Q., Luo, L., and Zheng, L. (2018). Lignins: Biosynthesis and biological functions in plants. *International Journal of Molecular Science* , **19** , 335.

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method.*Methods* , **25** , 402–408.

Llorente, F., López-Cobollo, R.M., Catalá, R., Martínez-Zapater, J.M. and Salinas, J. (2002) A novel cold-inducible gene from Arabidopsis, RCI3, encodes a peroxidase that constitutes a component for stress tolerance. *Plant Journal,* **32** , 13–24.

Ma, H., Chen, J., Zhang, Z., Ma, L., Yang, Z., Zhang, Q., *et al* . (2017) MAPK kinase 10.2 promotes disease resistance and drought tolerance by activating different MAPKs in rice. *Plant Journal,***92** , 557–570.

Manavalan, L.P., Prince, S.J., Musket, T.A., Chaky, J., Deshmukh, R., Vuong, T.D., *et al* . (2015) Identification of Novel QTL Governing Root Architectural Traits in an Interspecific Soybean Population.*PLoS One* , **10** , e0120490.

Miransari, M. and Smith, D.L. (2014) Plant hormones and seed germination. *Environmental and Experimental Botany,* **99** , 110–121.

Muhammad, T., Zhang, J., Ma, Y., Li, Y., Zhang, F., Zhang, Y. and Liang, Y. (2019) Overexpression of a mitogen-activated protein kinase*SlMAPK3* positively regulates tomato tolerance to cadmium and drought stress. *Molecules* , **24** , 556.

Munir, S., Liu, H., Xing, Y., Hussain, S., Ouyang, B., Zhang, Y., Li, H. and Ye, Z. (2016) Overexpression of calmodulin-like (*ShCML44*) stress-responsive gene from Solanum habrochaites enhances tolerance to multiple abiotic stresses. *Scientific Reports,* **6** , 31772.

Murray, M.G. and Thompson, W.F. (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research,* **8** , 4321–4326.

Muthusamy, M., Uma, S., Backiyarani, S., Saraswathi, M.S. and Chandrasekar, A. (2016) Transcriptomic changes of drought-tolerant and sensitive banana cultivars exposed to drought stress. *Frontiers in Plant Science,* **7** , 1609

Noctor, G., Mhamdi, A. and Foyer, C.H. (2014) The roles of reactive oxygen metabolism in drought: not so cut and dried. *Plant Physiology* , **164** , 1636-1648.

Pan, J., Li, Z., Wang, Q., Garrell, A.K., Liu, M., Guan, Y., Zhou, W. and Liu, W. (2018) Comparative proteomic investigation of drought responses in foxtail millet. *BMC Plant Biology* , **18** , 315.

Pandey, P., Irulappan, V., Bagavathiannan, M.V., and Senthil-Kumar, M. (2017). Impact of combined abiotic and biotic stresses on plant growth and avenues for crop improvement by exploiting physio-morphological traits. *Frontiers in Plant Science* , **8** , 537.

Peixoto-junior, R.F., Andrade, L.M. De, Brito, S., Nobile, P.M., Palma, A., Martins, B., *et al* . (2018) Overexpression of *ScMYBAS1* alternative splicing transcripts differentially impacts biomass accumulation and drought tolerance in rice transgenic plants. *PloS one* , **13** , e0207534.

Pou, A., Medrano, H., Flexas, J. and Tyerman, S.D. (2013) A putative role for TIP and PIP aquaporins in dynamics of leaf hydraulic and stomatal conductances in grapevine under water stress and re-watering. *Plant Cell and Environment,* **36** , 828–843.

Pourcel, L., Routaboul, J.M., Cheynier, V., Lepiniec, L. and Debeaujon, I. (2007) Flavonoid oxidation in plants: from biochemical properties to physiological functions. *Trends in Plant Science.* **12** , 29–36.

Qudeimat, E., Faltusz, A.M., Wheeler, G., Lang, D., Holtorf, H., Brownlee, C., Reski, R. and Frank, W., 2008. A PIIB-type Ca^{2+}-ATPase is essential for stress adaptation in Physcomitrella patens. *Proceedings of the National Academy of Sciences, U.S.A.* **105** , 19555–19560.

Raineri, J., Hartman, M.D., Chan, R.L., Iglesias, A.A. and Ribichich, K.F. (2016) A sunflower WRKY transcription factor stimulates the mobilization of seed-stored reserves during germination and post-germination growth. *Plant Cell Reports,* **35** , 1875–1890.

Akula, R. and Ravishankar, G.A. (2011) Influence of abiotic stress signals on secondary metabolites in plants. *Plant Signaling and Behaviour* , *6* , 1720-1731.

Rani, S., Schreinemachers, P. and Kuziyev, B. (2018) Mungbean as a catch crop for dryland systems in Pakistan and Uzbekistan : A situational analysis. *Cogent Food Agriculture,* **4** , 1499241.

Robinson, M.D., McCarthy, D.J. and Smyth, G.K. (2010) *edgeR: A Bioconductor package for differential expression analysis of digital gene expression data* . *Bioinformatics* , **26** , 139–140.

Salehin, M., Li, B., Tang, M., Katz, E., Song, L., Ecker, J.R., *et al* . (2019) Auxin-sensitive Aux/IAA proteins mediate drought tolerance in Arabidopsis by regulating glucosinolate levels. *Nature Communications,* **10** , 572305

Sanjari, S., Shirzadian-Khorramabad, R., Shobbar, Z.S. and Shahbazi, M. (2019) Systematic analysis of *NAC* transcription factors ' gene family and identification of post-flowering drought stress responsive members in sorghum. *Plant Cell Reports,* **38** , 361-376.

Sato, F. and Kumagai, H. (2013) Microbial production of isoquinoline alkaloids as plant secondary metabolites based on metabolic engineering research. *Proceedings of the Japan Academy Series B Physical Biological Sciences,* **89** , 165–182.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., *et al* . (2010) Genome sequence of the palaeopolyploid soybean. *Nature* , **463** , 178–183.

Semagn, K., Beyene, Y., Warburton, M.L., Tarekegne, A., Mugo, S., Meisel, B., *et al* . (2013) Meta-analyses of QTL for grain yield and anthesis silking interval in 18 maize populations evaluated under water-stressed and well-watered environments. *BMC Genomics* ,**14** , 313.

Severin, A.J., Woody, J.L., Bolon, Y., Joseph, B., Diers, B.W., Farmer, A.D., *et al* . (2010) RNA-Seq Atlas of Glycine max : A guide to the soybean transcriptome. *BMC Plant Biology* , **10** , 160.

Shao, H.B., Chu, L.Y., Shao, H.B., Chu, L.Y., Shao, M.A. and Zhao, C.X. (2008) Advances in functional regulation mechanisms of plant aquaporins: Their diversity, gene expression, localization, structure and roles

in plant soil-water relations. *Molecular Membrane Biology,* **25** , 179-191.

Sharmin, R.A., Bhuiyan, M.R., Lv, W., Yu, Z., Chang, F., Kong, J.,*et al* . (2020) RNA-Seq based transcriptomic analysis revealed genes associated with seed-flooding tolerance in wild soybean (*Glycine soja* Sieb. & Zucc.). *Environmental and Experimental Botany* , **171** , 103906.

Shi, W.Y., Du, Y.T., Ma, J., Min, D.H., Jin, L.G., Chen, J., *et al* (2018). *The WRKY* Transcription Factor *GmWRKY12* Confers Drought and Salt Tolerance in Soybean. *International Journal of Molecular Science,* **19** , 4087.

Shinozaki, K. and Yamaguchi-Shinozaki, K. (2007) Gene networks involved in drought stress response and tolerance. *Journal of Experimental Botany,* **58** , 221–227.

Shu, K., Zhou, W. and Yang, W. (2018) *APETALA* 2-domain-containing transcription factors: focusing on abscisic acid and gibberellins antagonism. *New Phytologist* , **217** , 977–983.

Singh, V.K., Jain, M. and Garg, R. (2015) Genome-wide analysis and expression profiling suggest diverse roles of *GH3* genes during development and abiotic stress responses in legumes. *Frontiers in Plant Science,* **5** , 1–13.

Sinha, A.K., Jaggi, M., Raghuram, B. and Tuteja, N. (2011) Mitogen-activated protein kinase signaling in plants under abiotic stress. *Plant Signaling and Behaviour,* **6** , 196–203.

Song, L., Prince, S., Valliyodan, B., Joshi, T., Maldonado, J. V, Wang, J., et al. (2016) Genome-wide transcriptome analysis of soybean primary root under varying water- deficit conditions. *BMC Genomics,***17** *,* 57.

Song, W.Y., Zhang, Z.B., Shao, H.B., Guo, X.L., Cao, H.X., Zhao, H.B., et al. (2008) Relationship between calcium decoding elements and plant abiotic-stress resistance. *International journal of Biological Sciences,* **4** , 116–125.

Stanke, M. and Waack, S. (2003) Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* , **19** , 215-225

Supek, F., Bošnjak, M., Škunca, N. and Šmuc, T. (2011) Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS One* ,**6** . e21800.

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., *et al* . (2014) *STRING v10:*Protein-protein interaction networks, integrated over the tree of life.*Nucleic Acids Research,* **43** , D447–D452.

Tang, Y., Bao, X., Zhi, Y., Wu, Q., Yin, X., Zeng, L., et al. (2019) *Overexpression of a MYB Family Gene , OsMYB6 , Increases Drought and Salinity Stress Tolerance in* transgenic rice. *Frontiers in Plant Sci* ence, *10* , 168.

Tenhaken, R. (2015) *Cell wall remodeling under abiotic stress .Frontiers in Plant Science, 5* , 771.

Thirumalaikumar, V.P., Devkar, V., Mehterov, N., Ali, S., Ozgur, R., Turkan, I., *et al* . (2018) *NAC* transcription factor*JUNGBRUNNEN1* enhances drought tolerance in tomato. *Plant Biotechnology Journal,* **16** , 354-366.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., et al. (2017) AgriGO v2.0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **45** , W122–W129.

Tiwari, S., Lata, C., Singh Chauhan, P., Prasad, V. and Prasad, M. (2017) A Functional Genomic Perspective on Drought Signalling and its Crosstalk with Phytohormone-mediated Signalling Pathways in Plants. *Current Genomics* , **18** , 469–482.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H. et al. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols,* **7** , 562–578.

Turlapati, P.V., Kim, K.W., Davin, L.B. and Lewis, N.G. (2011) The laccase multigene family in Arabidopsis thaliana: towards addressing the mystery of their gene function (s). *Planta* , **233** , 439-470.

Usadel, B., Poree, F., Nagel, A., Lohse, M., Czedik-Eysenberg, A.N.G.E.L.I.K.A. and Stitt, M. (2009) A guide to using MapMan to visualize and compare Omics data in plants: A case study in the crop species, Maize. *Plant Cell and Environment,* **32** , 1211–1229.

Valliyodan, B. and Nguyen, H.T. (2006) Understanding regulatory networks and engineering for enhanced drought tolerance in plants. *Current Opinions in Plant Biology,* **9** , 189–195.

Valliyodan, B., Ye, H., Song, L., Murphy, M., Shannon, J.G. and Nguyen, H.T., 2016. Genetic diversity and genomic strategies for improving drought and waterlogging tolerance in soybeans. *Journal of Experimental Bot* any, **68** , 1835-1849.

Van Der Straeten, D., Zhou, Z., Prinsen, E., Van Onckelen, H.A., and Van Montagu, M.C. (2001). A comparative molecular-physiological study of submergence response in lowland and deep water rice. *Plant Physiol* ogy, **125** , 955-968.

Vieira, R.D., TeKrony, D.M. and Egli, D.B. (1991) Effect of drought stress on soybean seed germination and vigor. *Journal of Seed Technology,***15** , 12-21.

Wang, C.T., Ru, J.N., Liu, Y.W., Yang, J.F., Li, M., Xu, Z.S. and Fu, J.D. (2018) The Maize *WRKY* transcription factor *ZmWRKY40*confers drought resistance in transgenic Arabidopsis. *International Journal of Molecular Sciences,***19** , 2580.

Wang, F., Kong, W., Wong, G., Fu, L., Peng, R., Li, Z. and Yao, Q. (2016) *AtMYB12* regulates flavonoids accumulation and abiotic stress tolerance in transgenic *Arabidopsis thaliana* .*Molecular Genetics and Genomics* , **291** , 1545–1559.

Wang, H. and Qin, F. (2017) Genome-wide association study reveals natural variations contributing to drought resistance in crops.*Frontiers in Plant Science.* , **8,** 1110.

Wang, J., Liu, S., Li, C., Wang, T., Zhang, P. and Chen, K. (2017)*PnLRR-RLK27,* a novel leucine-rich repeats receptor-like protein kinase from the Antarctic moss Pohlia nutans, positively regulates salinity and oxidation-stress tolerance. *PLoS One* , **12,**e0172869.

Wang, M., Yang, W.M. and Du, W.J. (2012). Construction of a molecular marker linkage map and its use for quantitative trait locus (QTLs) underlying drought tolerance at germination stage in soybean. *African Journal of Biotechnology,* **11,**12830-12838.

Wasaya, A., Zhang, X., Fang, Q. and Yan, Z. (2018) Root phenotyping for drought tolerance: A review. *Agronomy* **8** , 241.

Wei, S., Hu, W., Deng, X., Zhang, Y., Liu, X., Zhao, X., *et al* . (2014) A rice calcium-dependent protein kinase OsCPK9 positively regulates drought stress tolerance and spikelet fertility. *BMC Plant Biology.* , **14** , 133.

Wijewardana, C., Alsajri, F. A., Irby, J.T., Krutz, L.J., Golden, B., Henry, W.B., *et al* . (2019) Physiological assessment of water deficit in soybean using midday leaf water potential and spectral features. *Journal of Plant Interactions.* **14** , 533–543.

Wijewardana, C., Alsajri, F.A. and Reddy, K.R. (2019). Soybean seed germination response to *in vitro* osmotic stress. *Seed Technology.* **39** , 143–154.

Wilkins, K.A., Matthus, E., Swarbreck, S.M. and Davies, J.M. (2016) Calcium-mediated abiotic stress signaling in roots. *Frontiers in Plant Science* , **7** , 1296

Wu, C., Ding, X., Ding, Z., Tie, W., Yan, Y., Wang, Y., *et al* . (2019) The class III peroxidase (POD) gene family in cassava: Identification, phylogeny, duplication, and expression.*International Journal of Molecular Science* , **20** , 2730.

Wu, J., Jiang, Y., Liang, Y., Chen, L., Chen, W. and Cheng, B. (2019) Expression of the maize *MYB* transcription factor *ZmMYB3R*enhances drought and salt stress tolerance in transgenic plants.*Plant Physiology and Biochemistry* , **137** , 179–188.

Wu, L., Fang, Y., Yang, H. and Bai, L. (2019) Effects of drought-stress on seed germination and growth physiology of quinclorac-resistant*Echinochloa crusgalli* . *PLoS ONE* , 14, e0214480.

Wu, X., Qiao, Z., Liu, H., Acharya, B.R., Li, C. and Zhang, W., 2017. CML20, an Arabidopsis calmodulin-like protein, negatively regulates guard cell ABA signaling and drought stress tolerance. *Frontiers in Plant Science, 8* , 824.

Xie, X., He, Z., Chen, N., Tang, Z., Wang, Q. and Cai, Y., 2019. The Roles of Environmental Factors in Regulation of Oxidative Stress in Plant. *BioMed Research International,* 2019.

Xie, Z., Nolan, T.M., Jiang, H. and Yin, Y., 2019. AP2/ERF transcription factor regulatory networks in hormone and abiotic stress responses in Arabidopsis. *Frontiers in Plant Science,* **1** , 228.

Xuan, C.H.E.N., Zeng, Q.C., Lu, X.P., Yu, D.Q. and Li, W.Z., 2010. Characterization and expression analysis of four glycine-rich RNA-binding proteins involved in osmotic response in tobacco (*Nicotiana tabacum* cv. Xanthi). *Agricultural Sciences in China* , **9** , 1577–1587.

Yan, A., Wu, M., Yan, L., Hu, R., Ali, I. and Gan, Y., 2014. AtEXP2 is involved in seed germination and abiotic stress response in Arabidopsis. *PloS one* , **9** , e85208.

Yang, D.H., Kwak, K.J., Kim, M.K., Park, S.J., Yang, K., and Kang, H. (2014) Plant Science Expression of Arabidopsis glycine-rich RNA-binding protein *AtGRP2* or *AtGRP7* improves grain yield of rice (*Oryza sativa*) under drought stress conditions. *Plant Science* , **214** , 106–112.

Yang, L., Wang, C.C., Guo, W.D., Li, X.B., Lu, M., and Yu, C.L. (2006) Differential expression of cell wall related genes in the elongation zone of rice roots under water deficit. *Russ. J. Plant Physiology,* **53** , 390–395.

Yang, Y., Wang, Y., Jia, L., Yang, G., Xu, X., Zhai, H. et al (2018). Involvement of an ABI-like protein and a Ca^{2+}-ATPase in drought tolerance as revealed by transcript profiling of a sweetpotato somatic hybrid and its parents *Ipomoea batatas* (L.) Lam. and I. triloba L. *PloS One* , **13** , e0193193.

Yoon, G.M. and Kieber, J.J. (2013) 1-Aminocyclopropane-1-carboxylic acid as a signaling molecule in plants. *AoB Plants* , **5** , 1–6.

Zargar, S.M., Nagar, P., Deshmukh, R., Nazir, M., Wani, A.A., Masoodi, K.Z., et al. (2017) Aquaporins as potential drought tolerance inducing proteins: Towards instigating stress tolerance. *Journal of Proteomics* , **169** , 233–238.

Zhang, Q., Li, J., Zhang, W., Yan, S., Wang, R., Zhao, J., et al. (2012) The putative auxin efflux carrier OsPIN3t is involved in the drought stress response and drought tolerance. *Plant Journal,***72** , 805–816.

Zhang, X., Yang, G., Shi, R., Han, X., Qi, L., Wang, R., et al. (2013) Arabidopsis cysteine-rich receptor-like kinase 45 functions in the responses to abscisic acid and abiotic stresses. *Plant Physiology and Biochemistry* , **67** , 189–198.

Zhao, T., Aleem, M., Sharmin, R. A., 2017. Adaptation to Water Stress in Soybean: Morphology to Genetics. In: Plant, Abiotic Stress and Responses to Climate Change (Andjelkovic, V., eds), pp 33-68.

Croatia: IntechOpen,

Zhao, Y., Cheng, X., Liu, X., Wu, H., Bi, H. and Xu, H., 2018. The wheat MYB transcription factor TaMYB31 is involved in drought stress responses in Arabidopsis. *Frontiers in Plant Science* , **9** ,1426.

Zhou, F., Wu, G., Deng, W., Pu, Y., Wei, C., and Li, Y. (2007). Interaction of rice dwarf virus outer capsid P8 protein with rice glycolate oxidase mediates relocalization of P8. *FEBS Letters* , **581** , 34-40.

Table 1. Summary of Illumina sequencing and mapped reads data.

TF Class	**DTL**	**DTL**	**DSL**	**DSL**	**DSP**	**DSP**	**DTP**	**DTP**
	UP	DOWN	UP	DOWN	UP	DOWN	UP	DOWN
ARF	4	2	1	7	2	6	2	2
AP2	3	7	4	6	2	8	3	6
B3	2	4	1	5	2	4	3	3
bHLH	23	15	12	23	10	26	21	19
bZIP	5	4	2	5	3	8	3	7
BTB/POZ	8	6	3	11	5	9	8	4
C3H	4	5	2	5	3	9	4	4
Dof	5	7	2	8	3	10	4	9
ERF	30	18	15	28	9	30	12	21
GATA	4	4	0	3	1	3	0	4
HSF	3	12	0	16	1	12	1	12
LOB	2	6	1	5	1	8	5	4
MYB/ MYB-related	39	41	15	65	18	62	35	32
NAC	16	20	10	30	11	27	17	15
NF-YA	3	3	2	3	3	4	5	3
NF-YB	2	1	1	2	1	3	1	2
NF-YC	0	1	0	0	0	0	0	1
PHD	1	0	0	1	0	1	1	0
TCP	1	1	1	1	1	2	1	0
WRKY	28	10	11	38	16	33	27	19
ZF-HD	2	2	0	3	1	5	2	2
Total	**185**	**169**	**83**	**265**	**93**	**270**	**155**	**169**
	354		**348**		**363**		**324**	

Table 2 : Changes in the expression of hormone biosynthesis enzymes and regulating genes under drought stress.

Plant-hormones	**Enzymes involved in hormone biosynthesis**	**Transcriptionally responsive genes**
Auxin	TAA + ± /± +	Aux/IAA ± ± /± ±, GH3 ± ± /± ±, SAUR ± ±
	YUCCA ± ± /± ±	
Gibberellin	GA2Ox ± ± /± ±	GID1 + ± /± ±, DELLA - +- /+ -
	GA3Ox ± ± /± ±	
Abscisic acid	NCED - ± /- +-	HVA22 ± ± /± ±
	CCD ± ± / - +-	
Ethylene	ACO ± ± /± ±	EBF- - /- -, EIN2 ± ± /± ±, , EIN3 ± ± /± ±,
	ACS± ± /± ±	
Jasmonic acid	OPR ± ± /± ±	JAZ
Brassinosteroid	DET2 + + / + +	BRI1 + ± /+ +, BAK1 + ± /± ±

"+" = up -regulated genes; "–" = down-regulated genes; "±" = both up- and down-regulated genes exist. The four symbols under each gene represent directional changes in expression in: Drought tolerant line DTL, in drought sensitive line DSL, Drought sensitive Parent DSP and Drought tolerant Parent DTP, respectively.TAA = tryptophan aminotransferase of Arabidopsis; YUCCA = flavin-containing monooxygenase; Aux/IAA = auxin/indole acetic acid protein; GH3 = Gretchen Hagen 3; SAUR = small auxin-up RNA; ARF = auxin response factor; GA2OX = gibberellin 2 beta-dioxygenase; GA3OX = gibberellin 3 beta-dioxygenase GID1 = GA insensitive DWARF1;NCED = 9-cis-epoxycarotenoid dioxygenase; CCD = carotenoid cleavage dioxygenases; HVA22= *Hordeum Vulgare* ABA induced protein ; ABF = ABA-responsive element binding factors; ACO = 1-aminocyclopropane-1-carboxylate oxidase; ACS = aminocyclopropane-1-carboxylic acid synthase; EBF1/2 = EIN3-binding F-box protein 1/2; EIN2 = ethylene-insensitive protein 2; EIN3 = ethylene-insensitive protein 3; ERF = ethylene response factor.OPR3 = 12-oxophytodienoate reductase 3; JAZ1 = jasmonate-zim-domain protein 1; DET2= de-etiolated 2; BAK1= BRI1-associated receptor kinase 1; BRI; Brassinosteroid-Insensitive 1

Table 3: Details of differentially expressed TFs (upregulated and downregulated) identified by RNA-Seq under drought stress in soybean root tissues.

TF Class	**DTL**	**DTL**	**DSL**	**DSL**	**DSP**	**DSP**	**DTP**	**DTP**
	UP	DOWN	UP	DOWN	UP	DOWN	UP	DOWN
ARF	4	2	1	7	2	6	2	2
AP2	3	7	4	6	2	8	3	6
B3	2	4	1	5	2	4	3	3
bHLH	23	15	12	23	10	26	21	19
bZIP	5	4	2	5	3	8	3	7
BTB/POZ	8	6	3	11	5	9	8	4
C3H	4	5	2	5	3	9	4	4
Dof	5	7	2	8	3	10	4	9
ERF	30	18	15	28	9	30	12	21
GATA	4	4	0	3	1	3	0	4
HSF	3	12	0	16	1	12	1	12
LOB	2	6	1	5	1	8	5	4
MYB/ MYB-related	39	41	15	65	18	62	35	32
NAC	16	20	10	30	11	27	17	15
NF-YA	3	3	2	3	3	4	5	3
NF-YB	2	1	1	2	1	3	1	2
NF-YC	0	1	0	0	0	0	0	1
PHD	1	0	0	1	0	1	1	0
TCP	1	1	1	1	1	2	1	0
WRKY	28	10	11	38	16	33	27	19
ZF-HD	2	2	0	3	1	5	2	2
Total	**185**	**169**	**83**	**265**	**93**	**270**	**155**	**169**
	354		**348**		**363**		**324**	

Table 4: Differentially expressed genes (DEGs) related to cell wall metabolism, secondary metabolites, antioxidant activity, transcription factor activity, signaling pathway, aquaporin and hormonal signaling in drought tolerant Line (DTL), Drought sensitive Line (DSL), Drought Sensitive Parent (DSP) and Drought Tolerant Parent (DTP).

Trait Name	**Description**	**DTL**	**DTL**	**DSL**	**DSL**
Cell wall		UP	DOWN	UP	DOWN

Trait Name	Description	DTL	DTL	DSL	DSL
	PE	10	3	2	6
	XTH	15	6	5	12
	Expansins	15	3	8	11
	glycine-rich cell wall structural protein	1	0	0	1
	PLP	2	4	1	5
	MXAF	0	1	1	0
Shikimate acid pathway	SHK	1	0	0	0
	SDH	1	0	0	1
	CS	3	0	1	2
	ANPRT	1	1	0	2
	IGPS	3	0	1	2
	TS	4	0	2	2
Alkaloids biosynthetic pathway	STR	4	4	3	5
	DAT	2	0	1	1
Anthocyanin biosynthetic pathway	PAL	4	2	0	6
	TC4M	1	0	0	1
	4CL	2	2	1	3
	FLS	2	1	1	2
	F3H	1	0	0	1
	DFR	1	1	1	1
	ANR	2	0	0	2
	UFGT	2	2	1	3
Lignin biosynthetic pathway	COM	3	1	1	3
	CAD I	3	3	3	3
	SOH	1	2	0	3
	CA3M	4	2	2	4
	CCR	3	2	0	5
	LAC	8	2	3	7
Terpenoid biosynthetic pathway	HMGS	3	1	1	3
Antioxidants Activities	POD	25	9	9	28
	SOD	1	1	0	3
	Glutathione S-transferase	4	2	1	6
	Glutaredoxin (Grx)	1	0	0	2
	Thioredoxin	1	0	0	2
TF	Transcription Factors	185	169	83	265
Ca signaling	Calcium-binding proteins	4	1	1	4
	Calcium-dependent protein kinase	5	1	1	5
	Calcium ATPases	1	0	0	0
	Calmodulin-binding protein	2	1	0	3
	Calmodulin-like protein	3	1	0	3
	calmodulin-binding receptor	1	0	0	0
	Annexin	5	1	1	5
MAPK	Mitogen-activated protein kinases	4	1	0	2
RLK	Cysteine-rich -RLKs	4	1	1	4
	Leucine-rich repeat-RLKs	4	1	0	4
	Proline-rich-RLKs	2	0	0	2
Water transport	aquaporin	8	2	1	8
Hormones	ACC oxidase	9	2	1	7
	ACC synthase	4	1	1	5
	EIN3/ EIN2	4	2	1	5

Trait Name	Description	DTL	DTL	DSL	DSL
	Gibberellin 2-beta-dioxygenase	5	2	1	6
	Gibberellin 3-beta-dioxygenase	6	2	4	4
	Gibberellin 20 oxidase	4	2	2	4
	DELLA protein	2	1	0	1
	Abscisic acid 8'-hydroxylase	1	1	0	2
	NCED	0	2	1	1
	CCD	3	1	1	4
	Brassinosteroid Insensitive	2	0	1	2
	JA	2	0	1	0
	Auxin responsive protein IAA	8	3	1	7
	Auxin induced protein AUX	3	1	0	4
	auxin-responsive protein SAUR	4	2	1	4
	auxin efflux carrier component	3	1	1	7
	indole-3-acetic acid-amido synthetase GH3	3	1	1	5

PE; pectinesterase, XTH; xyloglucan endotransglucosylase PLP; Polygluctranose/ Pectate Lyase, MXAF; mannan-xylose-arabinose-fucose SDH, bifunctional 3-dehydroquinate dehydratase/shikimate dehydrogenase; SHK, shikimate kinase, CS, chorismate synthase; ANPRT, anthranilate phosphoribosyl transferase; IGPS, indole-3-glycerol phosphate synthase; TS, tryptophan synthase beta chain; STR, STRICTOSIDINE SYNTHASE; DAT,D-amino-acid transaminase phenylalanine ammonia-lyase;TC4M, trans-cinnamate 4-monooxygenase;4CL, 4-coumarate–CoA ligase; FLS, flavonol synthase/flavanone 3-hydroxylase; DFR, dihydroflavonol-4-reductase; ANR, anthocyanidin reductase; UFGT, anthocyanidin 3-O-glucosyltransferase; SOH, shikimate O-hydroxycinnamoyl transferase; COM, caffeoyl-CoA O-methyltransferase, CA3M, caffeic acid 3-O-methyltransferase; CAD, cinnamyl alcohol dehydrogenase; CCR, cinnamoyl-CoA reductase; LAC, laccase; HMGS, hydroxymethylglutaryl-CoA synthase; POD; Peroxidase, SOD; superoxide dismutase, ACO; 1-aminocyclopropane-1-carboxylate oxidase, ACS; aminocyclopropane-1-carboxylic acid synthase, EIN2; ethylene-insensitive protein 2, EIN3; ethylene-insensitive protein 3, NCED; 9-cis-epoxycarotenoid dioxygenase; CCD; carotenoid cleavage dioxygenases, JA; Jasmonic acid.

Table 5 : List of ten most possible candidate genes showing SNP polymorphism between drought-tolerant (DTL) and drought-sensitive (DSL) lines.

GENE	SNP POSITION	LOCATION	Changes in amino acid	DTL	DSL	Effect
Glyma.08G014200	1101347	Exon	Serine to Cysteine	C	G	Non-Synony
	1101776	Intron	Intron	A	G	
	1102052	Intron	Intron	C	C	
Glyma.08G029400	-	-	-	-	-	-
Glyma.08G042100	3338148	Exon	Lysine to glutamic acid	A	G	Non-synonyr
Glyma.08G045000	3568117	Exon	Tryptophan to tyrosine	T	G	Non-synonyr
	3568643	Exon	alanine to Valine	T	C	
	3569354	Exon	Aspartic acid to alanine	C	A	
	3569810	Exon	leucine to stop codon	T	A	
Glyma.08G091400	6865060	Exon	Cysteine to tyrosine	G	G	Non-synonyr
	6865062	Exon	isoleucine to Valine	A	A	
	6865084	Exon	Arginine to leucine	T	G	
	6865085	Exon	Histidine	T	C	
	6865136	Exon	phenylalanine to leucine	A	T	
	6865187	Exon	Valine	T	C	
	6865188	Exon	Glutamic acid to aspartic acid	G	C	

GENE	SNP POSITION	LOCATION	Changes in amino acid	DTL	DSL	Effect
	6865191	Exon	Arginine to glycine	A	G	
	6865226	Exon	Glycine	A	G	
	6865252	Exon	alanine to Valine	C	T	
	6865285	Exon	Valine to alanine	T	C	
	6865287	Exon	phenylalanine to isoleucine	T	A	
	6865289	Exon	glutamine to histidine	T	C	
	6865298	Exon	methionine to isoleucine	T	C	
	6865301	Exon	glutamic acid to aspartic acid	C	T	
	6865322	Exon	alanine	G	T	
	6865338	Exon	serine to proline	T	C	
	6865362	Intron	Intron	A	C	
	6865365	Intron	Intron	A	T	
	6865379	Intron	Intron	G	A	
	6865387	Intron	Intron	C	T	
Glyma.08G097300	7434448	Intron	Intron	T	C	
	7434721	Intron	Intron	A	G	
	7434937	Exon	leucine to phenylalanine	T	C	Non-synonyı
	7434944	Exon	methionine to leucine	T	C	
	7435000	Intron	Intron	T	C	
	7435009	Intron	Intron	T	A	
	7435010	Intron	Intron	A	C	
Glyma.08G162700	12738360	Exon	Serine to Phenylalanine	C	T	Non-Synony
Glyma.08G201700	16354631	Exon	Glycine to glutamine	A	C	Non-synonyı
	16354863	Exon	Threonine to Isoleucine	C	T	
Glyma.08G271600	35499076	Exon	Threonine to Isoleucine	C	T	Non-synonyı
	35499352	Exon	Valine to aspartic acid	T	T	
Glyma.19G184500	-	-	-	-	-	-

Hosted file

`Final all Figures.docx` available at https://authorea.com/users/308228/articles/439291-comprehensive-rna-seq-analysis-revealed-molecular-pathways-and-genes-associated-with-drought-tolerance-in-wild-soybean-glycine-soja-sieb-zucc