Iran. J. Biotechnol. October 2022;20(4): e3124 DOI: 10.30498/ijb.2022.299306.3124

Research Article

Co-Expression Network Analysis of Soybean Transcriptome Identify Hub Genes Under Saline-Alkaline and Water Deficit Stress

Rahele Ghanbari Moheb Seraj[1], Mahsa Mohammadi[2], Mehrdad Shahbazi[2], Masoud Tohidfar[*3]

[1]Department of Horticultural Sciences, Faculty of Agriculture and Natural Resources, University of Mohaghegh Ardabili, Ardabil, Iran
[2]Department of Plant Biotechnology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran
[3]Department of Plant Biotechnology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran

[*]*Corresponding author*: Masoud Tohidfar, Department of Plant Biotechnology, Faculty of Life Sciences and Biotechnology, Shahid Beheshti University, Tehran, Iran. Tel:+98-2122414413, Fax: +98-2122400067, E-mail: m_tohidfar@sbu.ac.ir

Background: Soybean is an important oilseed crop that its development and production are affected by environmental stresses (such as saline-alkaline and water deficit).

Objectives: This experiment was performed with the aim of identifying candidate genes in saline-alkaline stress and water-deficit stress conditions using transcriptome analysis and to investigate the expression of these genes under water deficit stress conditions using RTqPCR.

Materials and Methods: In this experiment, soybean transcriptome data under saline-alkaline and water-deficit stress were downloaded from the NCBI website, and then the co-expression modules were determined for them and the gene network was plotted for each module, and finally, the hub genes were identified. To compare the expression of genes in saline-alkaline and water deficit conditions, soybean plants were subjected to water deficit stress and their gene expression was determined using RTqPCR.

Results: The filtered (Log FC above +2 and below -2) genes of soybean were grouped under saline-alkaline stress in 15 modules and under water-deficit stress in 2 different modules. Within each module, the interaction of genes was identified using the gene network, then three genes of *ann11, cyp450* and *zfp* selected as hub genes. These hub genes are highly co-expression with other network genes, which not only display differential expression but also differential co-expression. The results of RT-PCR indicated that *cyp450* gene expression was not significantly different from the control, while *ann11* gene expression significantly increased under water deficit stress, but *zfp* gene expression decreased significantly under water deficit stress.

Conclusions: We identified three genes, *ann11, cyp450* and *zfp*, as hub genes. According to our results, *ann11* gene had a significant increase in expression under water deficit stress, which can indicate the importance of this gene under drought conditions. Therefore, according to the results of this experiment as well as other researchers, we introduce this gene as a key gene in water deficit tolerance and recommend its use in genetic engineering to increase the tolerance of other plants.

Keywords: Annexin, Cytochrome P450, $NaHCO_3$, R-Software, Water deficit, Zinc-finger protein

1. Background

Soybean is an important crop with several applications and is a source of protein and vegetable oil for human consumption as well as for the animal food industry. However, the quality and yield of soybean are affected by different stresses. Temperature, drought, salinity, and heavy metal stress affect the growth and yield of plants (1). These stresses are related to each other

Copyright *© 2022 The Author(s); Published by National Institute of Genetic Engineering and Biotechnology. This is an open access article, distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses /by-nc/4.0/) which permits others to copy and redistribute material just in noncommercial usages, provided the original work is properly cited.*

and may occur in the form of osmotic stress, ion distribution defect, and homeostasis in plant cells (2). The growth and production of plants are influenced by a response generated by a group of genes by changing their expression patterns. So, identifying responsive genes against abiotic stresses is necessary to understand the abiotic stress response mechanisms in crop plants. Plants can adapt to these stress conditions by the expression regulation of large sets of stress-related genes (3). Transcriptome data analysis will shed light on the transcriptional regulatory effect of abiotic and biotic stresses on gene expression in the cell. Gene expression microarray technology is an early developed but still utilized in biotechnological researches in which the genome-wide transcription information can be obtained for existing species with genome sequences or transcriptional loci (4). There are various methods for the identification of hub genes in transcriptomics data, one of which is the use of co-expression modules. Since most of the biological pathways involve coordinated regulation of dozens-to-hundreds of genes, known as co-expressed genes, so identification of these co-expressed genes or modules provides useful information for the characterization of key genes via systems-level approaches (5). Modules can be used to identify regulators, functional enrichment, and hub genes. These modules often demonstrate biological processes and can be a specific phenotype (6). Gene co-expression network analysis is one of the most powerful computational tools to identify the co-expressed gene modules and hub genes that play preponderant roles in constructed networks. Gene co-expression networks can be used for different aims, such as candidate gene prioritization, functional gene annotation, and identification of regulatory genes. Gene sets in the same co-expression modules are likely to be in the same functional category, participate in similar biological processes, or be part of the same pathway (7). Co-expression modules are often large and contain a large number of genes, so it is important to identify which gene in each module performs the main function. A widely used approach is to identify highly connected genes in a co-expression network (hub genes). Hubs are frequently more relevant to the functionality of networks than other nodes (8).

Many salt-inducible hub genes are also induced by drought stress and cold, which suggests the existence of similar mechanisms of stress responses (9). These hub genes can be used for the production of stress-tolerant plants either by genetic engineering or by molecular breeding. Several types of research on drought and salt stress using microarrays and transcriptome sequencing have identified stress-related genes involved in ion transport (annexin, zinc-finger proteins), cellular membrane integrity (annexin), scavenging of free oxygen radicals (cytochrome P450, zinc-finger proteins, annexin), and protecting macromolecules (cytochrome P450 and annexin) (10). Examination of transcriptome data and gene expression pattern in soybean in response to abiotic stresses can lead to the identification of genes effective in stress resistance in this plant.

2. Objectives

This experiment was performed with the aim of identifying candidate genes in saline-alkaline stress and water-deficit stress conditions using transcriptome analysis and to investigate the expression of these genes under water deficit stress conditions using RTqPCR. Identification of these genes helps us to improve and strengthen soybean plants under stress condition.

3. Material and Methods

3.1. Designation of Candidate Genes

Transcriptome profiles of wild soybean (*Glycine soja*) under $NaHCO_3$ treatment (Three week old Glycine soja seedling roots from 3 cm root apex were harvested in two independent biological replicates after 0, 0.5, 1, 3, 6, 12 and 24h treatment with 50mmol/L $NaHCO_3$ stress for RNA extraction and hybridization on Affymetrix microarrays.) and water deficits treatment (The expression of soybean genes in different root regions after 5 and 48 hours water stress treatment, and three biological replicates for each sample, using the Affymetrix Soybean GeneChip were examined.) were obtained from GEO DataSets, accession number GSE17883 (11) and GSE102749, respectively. The data with a P-value ≤ 0.01 were analyzed by the GEO2R program and the expression of treated genes compared to control genes was calculated and specified by logFC. Next, the files of all treatments were merged by R software to keep the genes common in all treatments. Then, the probes with logFC above +2 and below -2 were selected and saved. Gene names were determined by DAVID 6.8 program. The functional category of

genes was performed by ShinyGO v0.61 program (12). Gene groups' participation was determined by Venny 2.1 website (13). In the next step, data analysis was done using R, *GeneNet* package (14). First, co-expression gene modules were determined for all genes, within each module were placed genes with similar gene expression. Second, the co-expression diagrams and gene networks were plotted for each module. Finally, the networks were compared, and the genes with the high replication and interaction were selected.

3.2. Plant Material and Stress Treatment

Soybean seeds (*Glycine max*, cultivar Talar) were prepared from the Seed and Plant Improvement Institute, Karaj, Iran. The pot experiment was performed with two treatments (water deficit stress and control) and three replications. The soil composition of pots consisted of 1/3 clay, 1/3 sand, and 1/3 leaf composts. Three seeds were sown in pots (20 cm in diameter) at a depth of 3cm and then irrigated every two days. All pots were maintained at 25 °C with a photoperiod of 16 h and humidity of 50% (**Fig. 1**). After plant growth and establishment, water deficit was used to the stress treatment plants, in which the plants were placed under water deficit stress (no water) for 2 days, while the control treatment was irrigated as before (15). Then, the plants were separated from the *soil* surface and wrapped in sterile foil and placed in the -80 °C freezer.

3.3. Gene Expression Analysis

Hub genes were selected and their nucleotide sequences were received from the PLEXdb database and then BLAST was done at the NCBI website. Sequences that were highly similar were then selected to align with T-Coffee software. Primer design was performed from conserved regions (near the end of polyadenine). The PCR product (amplicon) length was between 200 to 250 bp. Homodimer, heterodimer, stem-loop, % GC, and Tm temperature of primers were calculated using Vector NTI (16) and Oligo7 software (17). The primer sequences of candidate and HKG genes along with their other characteristics are shown in **Table 1**. Total RNA was isolated from fresh roots using a total RNA kit (RB1001, RNA, Iran), and DNase I enzyme (RB125A, RNA, Iran) was used to remove genomic DNA contamination. Then, the first-strand cDNA was synthesized (RB125A, RNA, Iran) based on the manufacturer's instructions. qPCR with primers was performed to evaluate the performance and specificity as well as the setting of the primers. qPCR amplification was done with an Applied Biosystems™ StepOne™ qPCR System (Applied Biosystems, CA, USA) using SYBR® Green qPCR Master Mix (RB120, RNA, Iran). PCR reactions were provided in a final volume of 20 µL containing 1 µL of synthesized cDNA, 10 µL 2×SYBR, 2µL of each forward and reverce primers, and 7 µL distilled water. The reactions included a primary step

Figure 1. Overview of soybean pots in the greenhouse under controlled conditions

Table1. Primer sequences of candidate and housekeeping genes.

Gene accession number	Primer Name	Primer Sequence	Primer Length(bp)	Annealing temperature (°C)	Amplicon size	Primer efficiency
AFFX-r2-Bs-dap-3_at	*ann11*-F (5'→3')	AGAAATCGCTTACTCGTGTCG	21	56	~350bp	2.03
	ann11-R (3'→5')	GCAAGAAATCCTTGTAGCTC	20			
GmaAffx.93133.1.S1_s_at	*cyp450*-F (5'→3')	CTGTCTTCAGATGGATCAAG	20	56	~230bp	2.10
	cyp450-R (3'→5')	GAGCATGAGTATAGGTTGGAG	21			
Gma.17986.2.A1_at	*zfp*-F (5'→3')	TGCCTCACAAACTACAGC	18	54	~250bp	1.86
	zfp-R (3'→5')	GGACGCCTCATTGTTGGT	18			
AH001749.2	*18SrRNA*-F(5'→3')	ATGATAACTCGACGGATCGC	20	57	~170bp	1.98
	18SrRNA-R(3'→5')	CTTGGATGTGGTAGCCGTTT	20			

of 95 °C for 5 min, then 45 cycles at 95 °C (1 min) for denaturation, and 55 °C (1 min) for primer annealing (for all genes). Subsequently, the melting curves was plotted from range 55 °C to 95 °C in each reaction to check the specificity of the amplicons. Three biological replicates of samples (fresh roots) were employed for the RT-qPCR analysis, and three technical replicates were investigated for each biological replicate. $2^{-\Delta\Delta Ct}$ method was used for calculation of threshold cycle (Ct). Efficiency and Ct values were computed using the LinRegPCR quantitative PCR data analysis program (18), and relative expression values were analysed by REST method (19).

3.4. Statistical Analysis

Statistical analysis for logFC data of qPCR was done by R version 3.5.3 (14) and RStudio version 1.1.463 (20) softwares. Analysis of variance of the obtained data was performed in a completely randomized design. The mean comparison was done using the Duncan test method by R software and agricolae package at a significance level of 5%.

4. Results

4.1. Bioinformatics Analysis and Identification of Candidate Genes

4.1.1. Dataset of GSE17883 Under $NaHCO_3$ Treatment

The number of probes received from the GEO was 61170, which after filtering, their number was reduced to 321. These filtered genes (Differential Expression Genes (DEGs)) were investigated by ontological analysis and were classified based on function. The circular diagram of function classification indicated that these filtered genes are involved in various biological processes, such as response to stress, regulation of the metabolic process, cellular response to the stimulus, response to chemical, signaling, localization, developmental process, and reproduction. According to **Figure 2A**, a higher percentage of genes (69.28%) are devoted to the abiotic stress response. The plant response to environmental stress includes many genes and mechanisms, such as biochemical and molecular responses.

The results of the venny diagram also showed that out of 61170 genes received from NCBI, 321 genes had over-expression. Among the over-expressed genes, 19 genes were involved in response to stress and among the responsive genes, 3 hub genes were selected (**Fig. 3A**).

4.2. Dataset of GSE102749 Under Water Deficits Treatment

The number of 61169 probes were received from the GEO dataset and 7683 probes remained after filtering. Functional classification showed that the highest percentage of filtered genes (15%) were involved in stress response (**Fig. 2B**). In addition, 13% are involved in the organization of cellular components, 11% in the establishment of localization, 10% in the regulation of biological processes, 8% in the cellular response to stimuli, and the rest of the genes in other pathways listed in the pie chart (**Fig. 2B**). According to the results of Venny diagram, out of 61169 probes received from

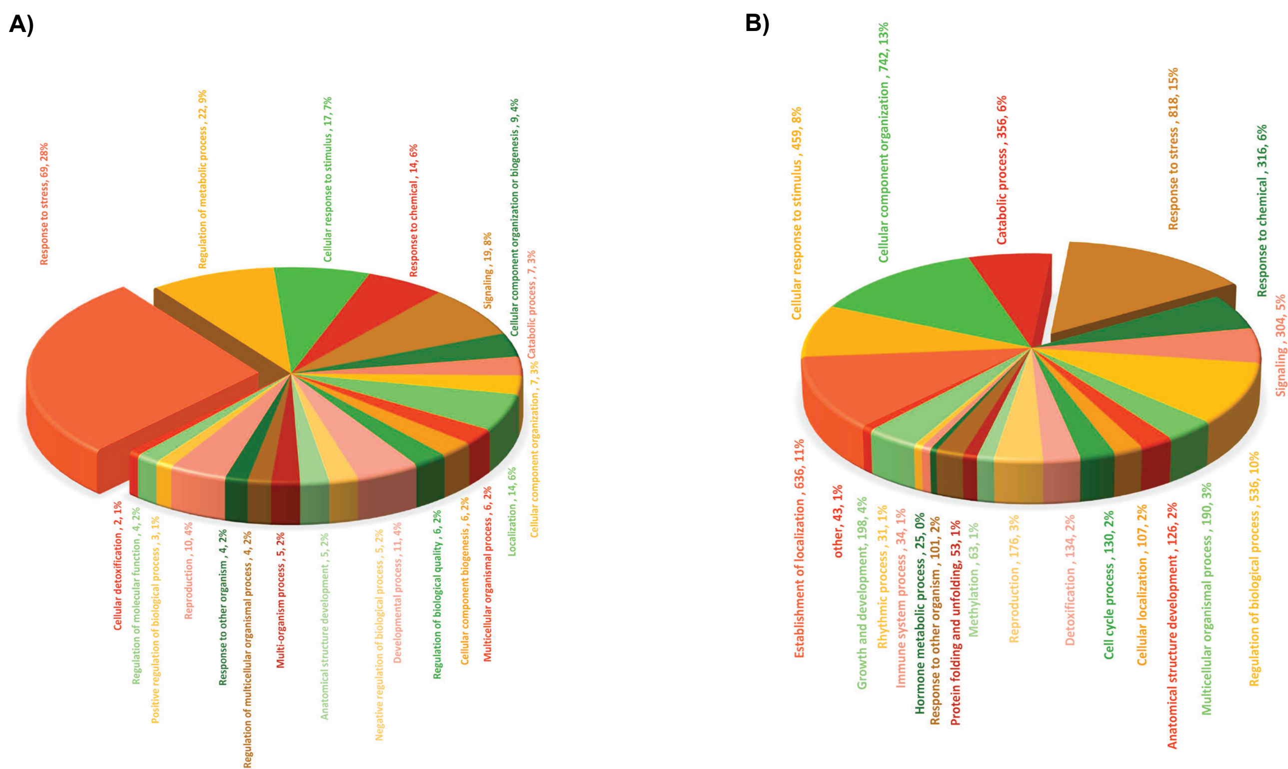


Figure 2. Ontological analysis of genes with differential expression in soybean, **A)** under $NaHCO_3$ treatment, **B)** under water deficits treatment

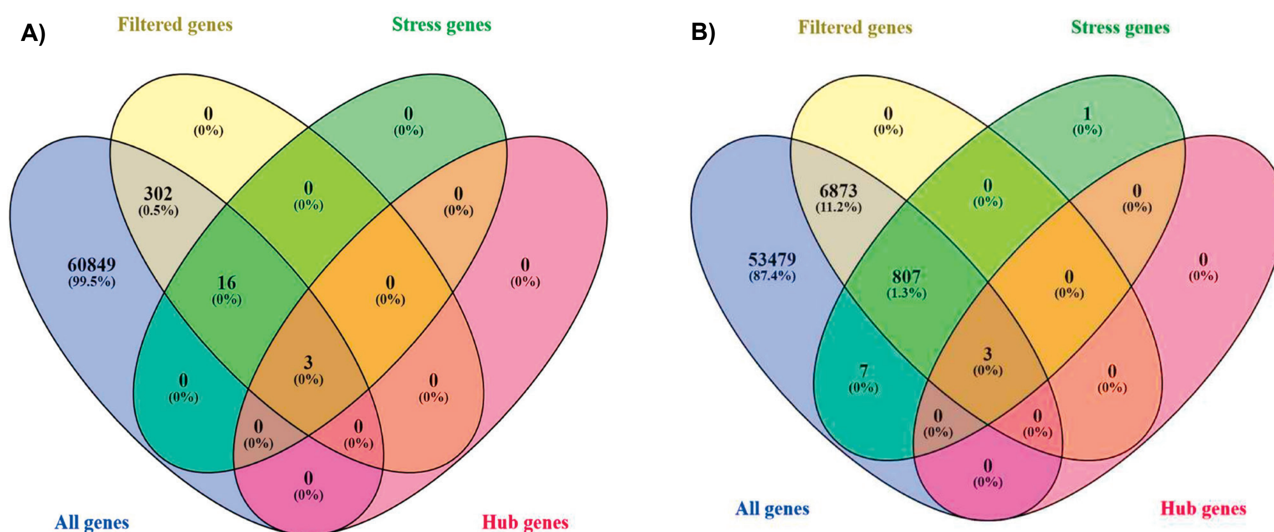


Figure 3. Venny diagram of several steps of soybean hub genes selection, **A)** under $NaHCO_3$ treatment, **B)** under water deficits treatment.

the GEO, 7683 genes had differential expression. Of these, 818 were involved in stress response and 3 of these stress-responsive genes were selected as hub genes and examined (**Fig. 3B**).

4.3. Identification Co-Expression Modules Under both $NaHCO_3$ and Water Deficit Treatments

The stress-responsive genes were grouped into different gene modules so that each module contained highly correlated genes. Pearson method and correlation coefficient of 0.01 were used to determine the correlation of genes. The mean of correlation in the genes inside each module under $NaHCO_3$ treatment is 98% and under water deficit treatment is 100%. Differential co-expression analysis can be used to recognize modules that behave differently under different conditions.

Generally, in $NaHCO_3$ treatment, the filtered genes were divided into 15 modules, of which 11 modules interacted with each other. Next, the co-expression diagram was plotted for the genes within these modules. As shown in Supplementary. 1A, the genes within each module showed similar expression under different treatments, but the pattern of gene expression was different in the several modules. In modules 1, 41, 53, 78, 241, 261, 304, 309, 327, 341, 376, and 381, approximately 3 h after saline-alkaline stress, gene expression was increased and then decreased drastically, but in gene modules 19 and 107, gene expression was strongly decreased in the early hours of stress but increased significantly at 24 hours after stress. Next, the gene name and expression of genes within each module were recorded in different treatments in Supplementary 2, in which the up-regulate and down-regulated genes are shown in green and red color, respectively. As is clear, the expression results of the genes in Supplementary 2 are in accordance with the results of the co-expression diagram (**Fig. 3B**) in related modules.

In water-deficit treatment, the filtered genes were placed in two modules that did not interact with each other. The co-expression diagram of these two modules showed that in module 1 the genes expression slightly decreased and in module 2 slightly increased (Supplementary. 1B). Next, the expression of genes within each module in different treatments was recorded in Supplementary 3. Up-regulated genes are green and down-regulated genes are red color. The results of Supplementary 3 correspond to the results of the co-expression diagram.

4.4. Co-Expression Gene Networks Analysis for Soybean Under both $NaHCO_3$ and Water Deficit Treatments

The gene network was plotted for the genes within each module (**Fig. 4**). The interaction between the genes was determined by two-way arrows and the non-interaction genes were deleted. Three genes were selected as candidate genes, including cytochrome P450 (*cyp450*: GmaAffx.93133.1.S1_s_at), zinc finger proteins (*zfp:* Gma.17986.2.A1_at), and annexin 11 (*ann11:* AFFX-r2-Bs-dap-3_at), which are represented by red, green, and orange colors, respectively. As can be seen in Supplementary 2 and 3, these three genes are visible with the mentioned colors and their number, full name of the gene and their exact identifier are also known. As the functional classification results have shown, these three genes are involved in abiotic stress tolerance. In the next step, the expression of three hub genes was examined under water deficit stress conditions.

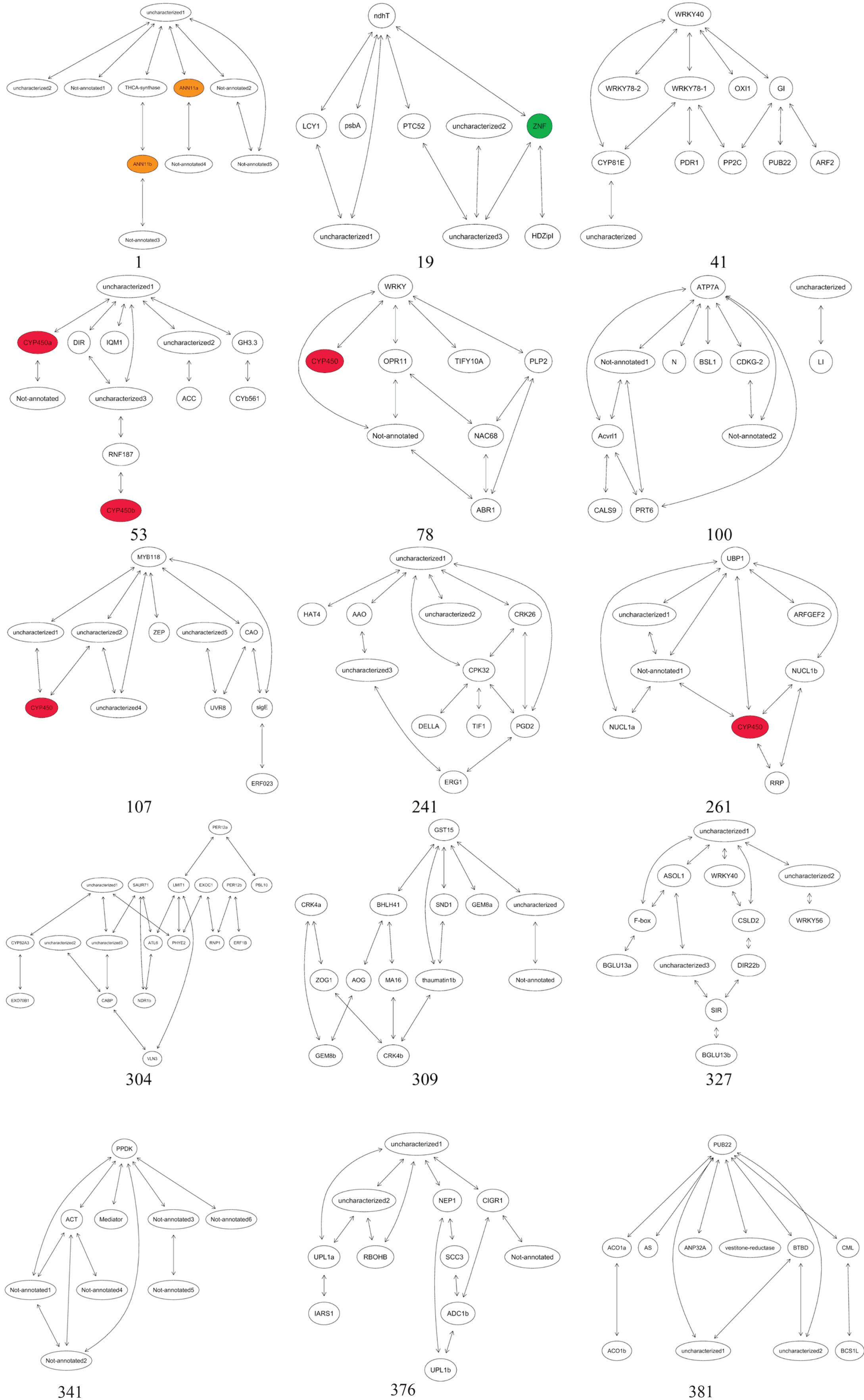
4.5. Gene Expression Analysis

Gene expression analysis using qPCR showed that *cyp450* gene expression under water deficit was not significantly different from control (LogFC=0.63). As shown in Supplementary 2, the expression of *cyp450* gene under conditions of water deficit decreased and is not consistent with the results of our experiment. According to Supplementary 3, *cyp450* gene expression increased significantly 6 hours after stress, and then showed a significant decrease 12 to 24 hours after stress. Unlike the *cyp450* gene, in our experiment the expression of the *zfp* gene was reduced (LogFC=0.0004). The expression of this gene was increased under water deficit treatment, but under $NaHCO_3$ treatment, *zfp* gene expression decreased up to 12h after stress and then increased. The *ann11* gene had an increase in expression in our experiment (LogFC=3.62). In the $NaHCO_3$ treatment, *ann11* gene expression increased 3h after stress and then decreased, but in the water deficit treatment, no significant difference was observed in the expression of *ann11* gene (**Fig. 5**).

5. Discussion

Soybean production is greatly influenced by abiotic stresses imposed by environmental factors such as drought, salt, and alkaline. In general, in this study, dataset GSE17883 under saline-alkaline treatment and dataset GSE102749 under treatment were examined.

A)



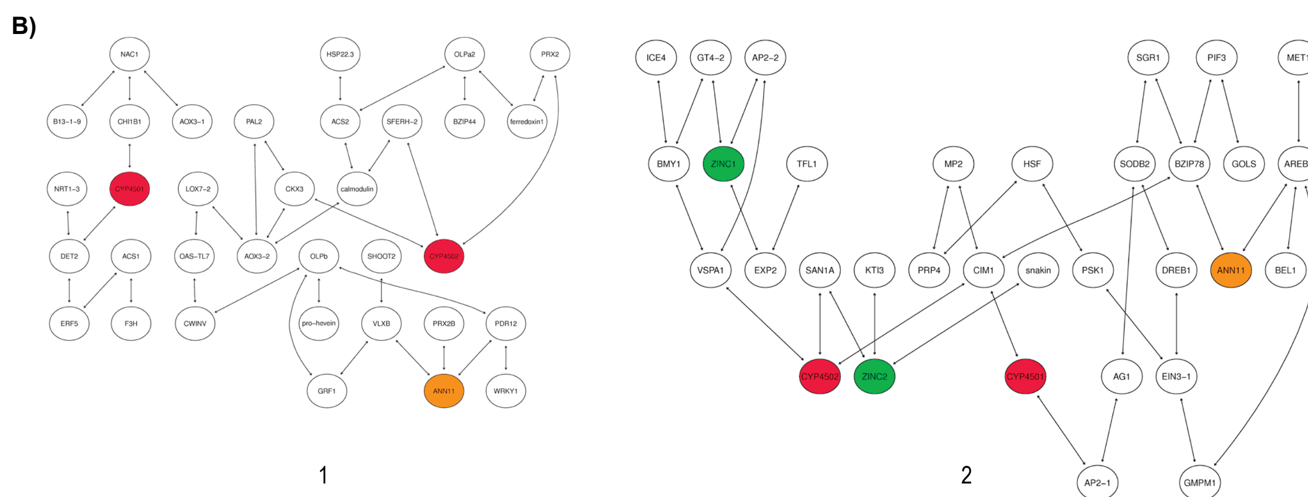


Figure 4. Gene networks of the module's genes. The numbers of each network is related to module group number. **A)** $NaHCO_3$ treatment, **B)** Water deficit treatment.

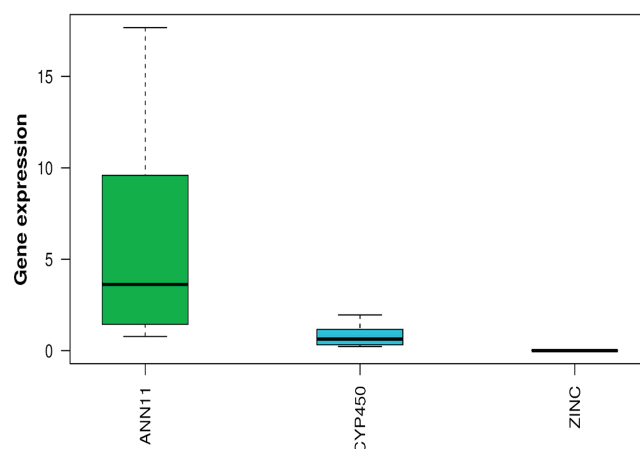


Figure 5. Expression ratio of three hub genes in soybean to the reference gene under drought stress.

Among the differential expression genes, three genes *cyp450*, *zfp* and *ann11* were selected as hub genes and their expression in soybean was studied under water deficit. The expression of these genes in soybean and other plants is widely studied. An ancient enzymatic system, is a diverse gene superfamily with extensive distribution in plants and animals (21). CYPs, as versatile catalysts, have vital roles in higher plants, in terms of compounds and metabolites biosynthesis, such as phytohormones, antioxidants, signal molecules, and structural polymers (22). The cytochrome P450 (*cyp450*) superfamily plays an important role in the plant's growth and development and plants protect from stresses through multiple biosynthetic and detoxification pathways (23).

In soybean, the *cyp450* gene family is including of 322 genes and 378 pseudogenes that their biological functions mostly have not been determined (24). Analysis of microarray data under $NaHCO_3$ treatment (GSE17883) showed that CYP450 gene expression increased up to 6 hours after saline-alkaline induction and then decreased sharply but the results of experiment GSE102749 showed that the expression of *cyp450* gene decreased under water-deficit stress. Our experimental results indicated that the expression of this gene under water deficit stress did not change significantly. Due to the high number of genes in the *cyp450* family, the genes examined in our experiment may be different from the genes in experiments GSE17883 and GSE102749, which will subsequently have different expressions.

Cytochrome P450 families in different plant species play an important role in increasing drought and salt tolerance (25). The plants have a relatively large number of *cyp450* genes and their extension in plants has created the largest family of enzymes in plant metabolism (26). According to supplementary 2 and 3, our results also indicated a high abundance of this gene and regulation of its expression under stress conditions. Wang *et al* (2016) reported that *ptcyp714a3*, a member of the cytochrome P450 gene family, is involved in salt resistance in rice (27). Li and Wei (2020) showed that under dehydration, 82 and 39 genes of *cyp450* family were identified in wheat and corn, respectively, which had significant expression (28). These genes may play a role in protecting plants from drought damage. In total, 1415 cytochrome gene sequences have been identified in plants of *Populus trichocarpa*, *Carica papaya*, *Arabidopsis thaliana*, *Oryza sativa*, *Physcomitrella patens* and *Vitis vinifera* (29). The expansion of this gene family in land plants has led to the production of the largest family of enzymes in plant metabolism (30). Numerous experiments have shown that *cyp450* plays an important role in plant adaptation to various abiotic stresses. In transgenic Arabidopsis plants, overexpression of *cyp709b3* genes increased salinity stress tolerance (31). In addition, overexpression of *pgcyp736b* gene increases salinity stress tolerance by reducing H_2O_2 accumulation, increasing carotenoid levels and expression of abscisic acid biosynthesis gene (25).

ZFP play essential roles in various organisms and their expression responses to different abiotic stresses (32). ZFPs are involved in ROS scavenging resulting from environmental stress and they adjusted the expression of stress-activated genes in plants via ROS signaling to resist salt, drought, or oxidative stress (33). In silico analysis in GSE17883 showed that *zfp* expression decreased up to 12 hours after induction of saline-alkaline stress but after 24 hours, *zfp* expression increased. Also, *zfp* expression in GSE102749 dataset under water-deficit treatment was increased. Our results indicated that, in water deficit conditions, gene expression was decreased after two days. Recent studies have shown that zinc finger proteins play important roles in plant development and also, plant tolerance to harmful stress (34). So far, many *zfp* genes have been cloned and studied in model plants such as Arabidopsis, wheat, soybeans and rice (33). These genes encode proteins that play a major role in growth and stress tolerance (35). Zinc finger proteins also improve salinity tolerance by increasing the concentration of osmotic regulators. *zfp252* and *zfp179* improve salinity tolerance in rice by increasing the content of free proline and soluble sugars, as well as by up-regulating various genes involved in the biosynthesis of osmotic substances, including *osdreb2a*, *osp5cs*, *osprot*, and *oslea3* (36). Zinc finger proteins increase the plant's drought resistance by increasing the level of osmotic regulators. Rice plants that overexpress *zfp252* have higher survival rates than wild-type and antisense-*zfp252* plants under drought stress (36). Excessive expression of *osmsr15* and *zfp3* in transgenic Arabidopsis plants increases drought tolerance by maintaining higher proline content, reducing electrolyte leakage, and increasing stress-responsive gene expression (37). These proteins also improve drought resistance by increasing the capacity to scavenge ROS. *zfp245* improves drought resistance in rice by increasing the activity of ROS scavenging enzymes, including *sod* and *pod*, and by increasing resistance to H_2O_2 (38).

ann11 are a multigene, multifunctional family of Ca^{2+}-dependent membrane-binding proteins found in both animal and plant cells, where they serve as important components of Ca^{2+} signaling pathways (39). They play a role in the organization of membrane protein networks and interaction with signaling pathway components and so they are linked to a wide range of cellular and developmental processes (40). Transcriptome analysis demonstrated that, under saline-alkaline stress, up to 3 hours after stress induction, the expression of *ann11* gene increased and then decreased, whereas, under water deficit treatment, no significant difference was observed in *ann11* expression. On the other hand, in our experimental conditions, the expression of this gene under drought stress was increased significantly compared to control. The role of annexins in abiotic stress tolerance has been extensively analyzed and multiple transgenic studies have indicated a positive impact of the expression of annexin genes on plant stress tolerance (41). Ijaz *et al*. (2017) reported that tomato plants with high *annsp2* expression increase drought tolerance and salinity. Transgenic tobacco plants that express the *annbj1* gene have a higher tolerance to abiotic stresses (42). Excessive expression of *ghann1* in cotton causes the plant to be more tolerant of environmental stresses (43). Aleem

et al. (2020) examined stress tolerance genes in wild soybean using RNA-seq method. Their results showed that many genes related to Ca^{2+} signaling in plants such as annexin had higher expression in drought tolerant genotypes than sensitive genotypes (44). Numerous experiments have shown the role of *ann* genes in the function of plant hormones such as abscisic acid, gibberellic acid, jasmonic acid and salicylic acid (43). ABA regulates *ann* expression in a wide range of plant species (45). Drought and salinity stress significantly increases the level of abscisic acid, which in turn causes the expression of many genes involved in stress (46). Researchers' reports indicate that *gsrlck*, *gmgras37* and *gsgst* expression was induced by ABA, salt and drought stresses (47). According to Sun *et al*. (2013), overexpression of *gsrlck* reduces ABA sensitivity and alters the expression levels of ABA-responsive genes. Gene expression is one of the sensitive analyses that may be different in two plants under the same conditions. These changes may be due to environmental conditions including humidity, light, soil type, etc. In some conditions, high stress may be applied to the plant and in some conditions, low stress. In addition to environmental conditions, gene expression levels correlate with multiple aspects of gene sequence and structure including synonymous codon usage (48).

Most of the analysis of this study was performed using the R program. R, a free software environment, is a unified set of software facilities for data manipulation, calculation, and visualization. Therefore, data analysis with this method is free, fast, and accurate, besides, it provides many options for subsequent analysis.

6. Conclusion

In this study, we used the co-expression analysis method to recognize stress-responsive genes. Gene co-expression network analysis separated DEG sets into numerous modules that display high co-expression and determined stress-related hub genes with a high connection. Our hub genes in this experiment are *ann11*, *cyp450*, and *zfp* genes. Our results and other researchers showed that these three genes play a key role in abiotic stress tolerance. In this study, we compared the expression of these genes in saline-alkaline and drought conditions. In general, the results showed that the gene expression of *cyp450* gene did not differ significantly from control, while *ann11* gene increased under water deficit stress, but *zfp* gene expression decreased significantly under water deficit stress.

References

1. Vaughan MM, Block A, Christensen SA, Allen LH, Schmelz EA. The effects of climate change associated abiotic stresses on maize phytochemical defenses. *Phytochem Rev*. 2018;**17**(1):37-49. doi: 10.1007/s11101-017-9508-2
2. Aydin A, Kurt F, Hürkan K. Key aromatic amino acid players in soybean (Glycine max) genome under drought and salt stresses. *Biocatal Agric Biotechnol*. 2021;**35**:102094. doi: 10.1016/j.scienta.2019.02.044
3. Wang H, Wang H, Shao H, Tang X. Recent advances in utilizing transcription factors to improve plant abiotic stress tolerance by transgenic technology. *Front Plant Sci*. 2016;**7**:67. doi: 10.3389/fpls.2016.00067
4. Almeida-Silva F, Moharana KC, Venancio TM. The state of the art in soybean transcriptomics resources and gene coexpression networks. *In Silico Plants*. 2021;**3**(1):diab005. doi: 10.1093/insilicoplants/diab005
5. Lv Y, Xu L, Dossa K, Zhou K, Zhu M, Xie H, *et al*. Identification of putative drought-responsive genes in rice using gene co-expression analysis. *Bioinformation*. 2019;**15**(7):480. doi: 10.6026/97320630015480
6. Langfelder P, Mischel PS, Horvath S. When is hub gene selection better than standard meta-analysis? *PLoS One*. 2013;8(4):e61505. doi: 10.1371/journal.pone.0061505
7. Ma S, Gong Q, Bohnert HJ. An Arabidopsis gene network based on the graphical Gaussian model. *Genome Res*. 2007;**17**(11):1614-1625. doi: 10.1101/gr.6911207
8. Albert R, Jeong H, Barabási A-L. Error and attack tolerance of complex networks. *Nature*. 2000;**406**(6794):378-382. doi: 10.1038/35019019
9. Ciarmiello L, Woodrow P, Fuggi A. Plant genes for abiotic stress, abiotic stress in plants–mechanisms and adaptations, Prof. Arun Shanker (Ed.), ISBN: 978-953-307-394-1, *INTECH*. 2011. doi: 10.5772/22465
10. Veith A, Moorthy B. Role of cytochrome P450s in the generation and metabolism of reactive oxygen species. *Curr Opin Toxicol*. 2018;**7**:44-51. doi: 10.1016/j.cotox.2017.10.003
11. Ge Y, Li Y, Zhu Y-M, Bai X, Lv D-K, Guo D, *et al*. Global transcriptome profiling of wild soybean (Glycine soja) roots under $NaHCO_3$ treatment. *BMC Plant Biol*. 2010;**10**(1):1-14. doi: 10.1186/1471-2229-10-153
12. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. *Bioinformatics*. 2020;**36**(8):2628-2629. doi: 10.1093/bioinformatics/btz931
13. Oliveros J. Venny. An interactive tool for comparing lists with Venn's diagrams. 2016;2007-2015. doi: 10.1371/journal.pone.0154315
14. Team RC. R Foundation for Statistical Computing; Vienna, Austria: 2018. R: A language and environment for statistical computing. 2017.
15. Prince SJ, Murphy M, Mutava RN, Durnell LA, Valliyodan B, Shannon JG, *et al*. Root xylem plasticity to improve water use and yield in water-stressed soybean. *J Exp Bot*. 2017;**68**(8):2027-2036. doi: 10.1093/jxb/erw472
16. Lu G, Moriyama EN. Vector NTI, a balanced all-in-one sequence analysis suite. *Brief Bioinform*. 2004;**5**(4):378-388. doi: 10.1093/bib/5.4.378

17. Rychlik W. OLIGO 7 primer analysis software. *PCR primer design*. 2007:35-59. doi: 10.1007/978-1-59745-528-2_2
18. Ruijter J, Ramakers C, Hoogaars W, Karlen Y, Bakker O, Van den Hoff M, *et al*. Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res*. 2009;**37**(6):e45-e. doi: 10.1093/nar/gkp045
19. Pfaffl MW, Horgan GW, Dempfle L. Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Res*. 2002;**30**(9):e36-e. doi: 10.1093/nar/30.9.e36
20. Allaire J. RStudio: integrated development environment for R. Boston, MA. 2012;**770**(394):165-171.
21. Shiota N, Kodama S, Inui H, Ohkawa H. Expression of human cytochromes P450 1A1 and P450 1A2 as fused enzymes with yeast NADPH-cytochrome P450 oxidoreductase in transgenic tobacco plants. *Biosci Biotechnol Biochem*. 2000;**64**(10):2025-2033. doi: 10.1271/bbb.64.2025
22. Renault H, Bassard J-E, Hamberger B, Werck-Reichhart D. Cytochrome P450-mediated metabolic engineering: current progress and future challenges. *Curr Opin Plant Biol*. 2014;**19**:27-34. doi: 10.1016/j.pbi.2014.03.004
23. Li Z-Y, Xu Z-S, He G-Y, Yang G-X, Chen M, Li L-C, *et al*. Overexpression of soybean GmCBL1 enhances abiotic stress tolerance and promotes hypocotyl elongation in Arabidopsis. *Biochem Biophys Res Commun*. 2012;**427**(4):731-736. doi: 10.1016/j.bbrc.2012.09.128
24. Guttikonda SK, Trupti J, Bisht NC, Chen H, An Y-QC, Pandey S, *et al*. Whole genome co-expression analysis of soybean cytochrome P450 genes identifies nodulation-specific P450 monooxygenases. *BMC Plant Biol*. 2010;**10**(1):1-19. doi: 10.1186/1471-2229-10-243
25. Balusamy SR, Rahimi S, Yang D-C. Characterization of squalene-induced PgCYP736B involved in salt tolerance by modulating key genes of abscisic acid biosynthesis. *Int J Biol Macromol*. 2019;**121**:796-805. doi: 10.1186/1471-2229-10-243
26. Shi Y, Wang H, Liu Z, Wu S, Yang Y, Feyereisen R, *et al*. Phylogenetic and functional characterization of ten P450 genes from the CYP6AE subfamily of Helicoverpa armigera involved in xenobiotic metabolism. *Insect Biochem Mol*. 2018;**93**:79-91. doi: 10.1016/j.ibmb.2017.12.006
27. Wang C, Yang Y, Wang H, Ran X, Li B, Zhang J, *et al*. Ectopic expression of a cytochrome P450 monooxygenase gene PtCYP714A3 from Populus trichocarpa reduces shoot growth and improves tolerance to salt stress in transgenic rice. *Plant Biotechnol J*. 2016;**14**(9):1838-1851. doi: 10. 1111/pbi.12544
28. Li Y, Wei K. Comparative functional genomics analysis of cytochrome P450 gene superfamily in wheat and maize. *BMC Plant Biol*. 2020;**20**(1):1-22. doi: 10.1186/s12870-020-2288-7
29. Nelson DR, Ming R, Alam M, Schuler MA. Comparison of cytochrome P450 genes from six plant genomes. *Trop Plant Biol*. 2008;**1**(3):216-235. doi: 10.1007/s12042-008-9022-1
30. Nelson D, Werck-Reichhart D. A P450-centric view of plant evolution. *Plant J*. 2011;**66**(1):194-211. doi: 10.1111/j.1365-313X.2011.04529.x
31. Mao G, Seebeck T, Schrenker D, Yu O. CYP709B3, a cytochrome P450 monooxygenase gene involved in salt tolerance in Arabidopsis thaliana. *BMC Plant Biol*. 2013; **13**(1):1-13. doi: 10.1186/1471-2229-13-169
32. Luo X, Bai X, Zhu D, Li Y, Ji W, Cai H, *et al*. GsZFP1, a new Cys2/His2-type zinc-finger protein, is a positive regulator of plant tolerance to cold and drought stress. *Planta*. 2012;**235**(6):1141-1155. doi: 10.1007/s00425-011-1563-0
33. Zhang H, Liu Y, Wen F, Yao D, Wang L, Guo J, *et al*. A novel rice C2H2-type zinc finger protein, ZFP36, is a key player involved in abscisic acid-induced antioxidant defence and oxidative stress tolerance in rice. *J Exp Bot*. 2014;**65**(20):5795-5809. doi: 10.1093/jxb/eru313
34. Wang Y, Jiang L, Chen J, Tao L, An Y, Cai H, *et al*. Overexpression of the alfalfa WRKY11 gene enhances salt tolerance in soybean. *PLoS One*. 2018;**13**(2):e0192382. doi: 10.1371/journal.pone.0192382
35. Cao H, Huang P, Zhang L, Shi Y, Sun D, Yan Y, *et al*. Characterization of 47 Cys2 -His2 zinc finger proteins required for the development and pathogenicity of the rice blast fungus Magnaporthe oryzae. *New Phytol*. 2016;**211**(3):1035-1051. doi: 10.1111/nph.13948
36. Xu DQ, Huang J, Guo SQ, Yang X, Bao YM, Tang HJ, *et al*. Overexpression of a TFIIIA-type zinc finger protein gene ZFP252 enhances drought and salt tolerance in rice (Oryza sativa L.). *FEBS Lett*. 2008;**582**(7):1037-1043. doi: 10.1016/j.febslet.2008.02.052
37. Han G, Lu C, Guo J, Qiao Z, Sui N, Qiu N, *et al*. C2H2 Zinc Finger Proteins: Master Regulators of Abiotic Stress Responses in Plants. *Front Plant Sci*. 2020;**11**:115. doi: 10.3389/fpls.2020.00115
38. Huang J, Sun SJ, Xu DQ, Yang X, Bao YM, Wang ZF, *et al*. Increased tolerance of rice to cold, drought and oxidative stresses mediated by the overexpression of a gene that encodes the zinc finger protein ZFP245. *Biochem Biophys Res Commun*. 2009;**389**(3):556-561. doi: 10.1016/j.bbrc.2009.09.032
39. Mortimer JC, Laohavisit A, Macpherson N, Webb A, Brownlee C, Battey NH, *et al*. Annexins: multifunctional components of growth and adaptation. *J Exp Bot*. 2008;**59**(3):533-544. doi: 10.1093/jxb/erm344
40. Gerke V, Creutz CE, Moss SE. Annexins: linking Ca^{2+} signalling to membrane dynamics. *Nat Rev Mol Cell Biol*. 2005;**6**(6):449-461. doi: 10.1038/nrm1661
41. Ijaz R, Ejaz J, Gao S, Liu T, Imtiaz M, Ye Z, *et al*. Overexpression of annexin gene AnnSp2, enhances drought and salt tolerance through modulation of ABA synthesis and scavenging ROS in tomato. *Sci Rep*. 2017;**7**(1):12087. doi: 10.1038/s41598-017-11168-2
42. Jami SK, Clark GB, Turlapati SA, Handley C, Roux SJ, Kirti PB. Ectopic expression of an annexin from Brassica juncea confers tolerance to abiotic and biotic stress treatments in transgenic tobacco. *Plant Physiol Biochem*. 2008;**46**(12):1019-1030. doi: 10.1016/j.plaphy.2008.07.006
43. Zhang F, Li S, Yang S, Wang L, Guo W. Overexpression of a cotton annexin gene, GhAnn1, enhances drought and salt stress tolerance in transgenic cotton. *Plant Mol Biol*. 2015;**87**(1-2):47-67. doi: 10.1007/s11103-014-0260-3
44. Aleem M, Raza MM, Haider MS, Atif RM, Ali Z, Bhat JA, *et al*. Comprehensive RNA-seq analysis revealed molecular pathways and genes associated with drought tolerance in wild soybean (Glycine soja Sieb. and Zucc.). *Physiol Plant*. 2021;**172**(2):707-732. doi: 10.1111/ppl.13219
45. Lee S, Lee EJ, Yang EJ, Lee JE, Park AR, Song WH, *et al*. Proteomic identification of annexins, calcium-dependent

membrane binding proteins that mediate osmotic stress and abscisic acid signal transduction in Arabidopsis. *Plant Cell*. 2004;**16**(6):1378-1391. doi: 10.1105/tpc.021683
46. Tuteja N. Mechanisms of high salinity tolerance in plants. *Methods Enzymol*. 2007;**428**:419-438. doi: 10.1016/S0076-6879(07)28024-3
47. Sun X, Sun M, Luo X, Ding X, Cai H, Bai X, *et al*. A Glycine soja ABA-responsive receptor-like cytoplasmic kinase, GsRLCK, positively controls plant tolerance to salt and drought stresses. *Planta*. 2013;**237**(6):1527-1545. doi: 10.1007/s00425-013-1864-6
48. Williford A, Demuth JP. Gene expression levels are correlated with synonymous codon usage, amino acid composition, and gene architecture in the red flour beetle, Tribolium castaneum. *Mol Biol Evol*. 2012;**29**(12):3755-3766. doi: 10.1093/molbev/mss184