http://www.pjbs.org

PJBS

ISSN 1028-8880

Pakistan Journal of Biological Sciences

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Pakistan Journal of Biological Sciences 17 (1): 80-85, 2014
ISSN 1028-8880 / DOI: 10.3923/pjbs.2014.80.85

Bioinformatical Evaluation of Desiccation-responsive *rd29A* Gene in *Arabidopsis thaliana*

Mostafa Modarresi and Alireza Eyvazi
Department of Biology, Shahre-Rey Branch, Islamic Azad University,
Tehran, Islamic Republic of Iran

Abstract: In *Arabidopsis thaliana* expression of two Desiccation-Responsive *rd29* genes are intensity induced under drought, salinity and temperature (heat and cold) conditions. These two genes called, *rd29A* and *rd29B* are located in 8-9 Kb of Arabidopsis genome. Also, transgenic plants containing these genes have more tolerant toward environmental stresses. Since *rd29A* has an important role in various non-biological stresses, in this study genetic and promoter analysis of this gene was performed. *rd29A* nucleotide sequence was obtained from NCBI gene bank, for bioinformatic analysis used the Map viewer, Plant care and Genevestigator software. In Arabidopsis *rd29A* has 8048 nucleotides and encode a linear mRNA with 2133bp. It located at AT5G52310 locus on chromosome 5 between nucleotides 21241898 and 21242457 and contains 4 exons and 3 introns. Promoter region analysis to determine the gene regulatory elements performed with using 1500 bp of the 5 'UTR sequence and the results indicated that the existence of diverse regulatory elements such as ABRE, DRE, Box I, C-repeat/DRE and TCA-element which represent involvement of this gene in different physiological pathways in plants. In addition the TATA-Box was identified as Core promoter in -30 position. Also, the pattern of gene expression was examined by Microarrays data and it was found that the expression levels were variable in different developmental stages, so that the lowest and highest expressions were in germination and mature stages. It seems that using of this gene promoter in expression cassettes can raise tolerance of transgenic plants to environmental stresses.

Key words: *Arabidopsis thaliana*, bioinformatics, *rd29A* gene

INTRODUCTION

Abiotic stresses conditions such as drought, high salinity and temperature stress (low or high) are the major environmental factors to limit growth and agricultural productivity (Modarresi *et al*., 2012, 2013). Plants have number of physiological and biochemical responses which can reduce stress corresponding damage. Molecular studies have been shown that the expression of nine genes associated with these stresses and responsive to dehydration induced (Yamaguchi-Shinozaki and Shinozaki, 1994). The *rd29* (Responsive to Desiccation 29) is one of the nine RD genes of *Arabidopsis thaliana*, which isolated from this plant under drought stress condition. The *rd29* is induced by dehydration stress very quickly (within 20 min) and strongly (within 3 hours), and this induction is a two-phase process Extensive restriction analysis has revealed that *rd29* includes two genes, *rd29A* and *rd29B* (Yamaguchi-Shinozaki and Shinozaki 1993b). *rd29* sequences are used as markers to monitor stress-response pathways in plants because its exquisitely sensitive to various abiotic stress and differently change in the response to drought, cold and freezing, salt stress, or exposure to ABA. Due to the unique elemental combinations, *rd29A* and *rd29B* are differentially induced under abiotic stress conditions; the promoter of *rd29A* was found to be more responsive to drought and temperature stresses, since the promoter of *rd29B* was highly responsive to salt stress (Msanne *et al*., 2011).

Arabidopsis thaliana is a model plant growing in Europe and Asia. *Arabidopsis thaliana* has a relatively small genome with approximately 135 Mega base pairs (Mbp) that widely used for genetically studies since it has short life cycle and small size (Miller *et al*., 2010).

Transgenic plants containing these genes (*rd29A* and *rd29B*) also are more resistant to environmental stresses. Transformation and expression of the *rd29A*

Corresponding Author: Mostafa Modarresi, Department of Biology, Shahre-Rey Branch, Islamic Azad University, Tehran, Islamic Republic of Iran

promoter caused improving plant tolerance to drouht and salinity stresses, thus in this research investigated bioinformatics analysis of *rd29A* gene in *Arabidopsis thaliana* as a model plant.

MATERIALS AND METHODS

Gene location and locus number of *rd29A* gene from *Arabidopsis thaliana* obtained from PHYTOZOME site available at (www.phytozome.net). For determination gene regulatory elements, used 1500 bp of the 5' UTR for promoter region analysis by Place (http://www.dna.affrc.go.jp/PLACE/) and Plant care (http://bioinformatics.psb.ugent.be/webtools/plantcare/html) programs, then motifs extracting in the promoter region and examined their function. Also gene expression pattern analyzed by Genevestigator program (www.genevestigator.com) and Microarray data. *rd29A* protein domains determined by SMART (www.smart.embl-heidelberg.de) and STRING (http://string90.embl.de/), then genetic network designed. Protein Haplotype diversity analyzed by DnaSP software. Phylogenetic relationships of this protein performed by MEGA 5.0 software and the Neighbor-joining method.

RESULTS

The *rd29A* nucleotide sequence obtained from NCBI gene bank, also used Map viewer, Plant care and Genevestigator software for bioinformatics analysis. *Arabidopsis thaliana rd29A* gene has 8048 nucleotides, that encoding 2133 bp linear mRNA in locus AT5G52310 on chromosome 5 between nucleotides 21241898 and 21242457. This gene has 4 exons and 3 introns. Using 1500 bp of the 5' UTR promoter region analysis to determine the gene regulatory elements performed and the results indicate that existence of diverse regulatory elements such as ABRE (cis-acting element in the response to stress of various non-biological), DRE (Effective in response to drought, cold and salinity with (TACCGACAT) 9 bp sequence), Box I (element responds to light), C-repeat/DRE (in response to cold and drought stress) and TCA-element (in response to salicylic acid) which represents involvement of these genes in different physiological pathways in plants (Table 1). In addition identified TATA Box as Core promoter in -30 position. ProtParam (http://web.expasy.org/protparam/) analysis indicated that *rd29A* is a protein with 710 amino acids and 77.885 KDa molecular weight and pI = 4.45. Total number of negatively (Asp+Glu) and positively (Arg+Lys) charged residues are 168 and 78 respectively. This protein has 10650 atoms with $C_{3342}H_{5172}N_{912}O_{1215}S_9$ formula. The Instability Index (II) is computed to be 53.32; this classifies the protein as unstable. Also *rd29A* aliphatic index is 48.82 and Grand average of hydropathicity (GRAVY) is -1.087. Uniprot (http://www.uniprot.org/) data indicated that this protein has repeated sequence of 2 X 14 AA repeats of P-[MV]-G-F-G-[DS]-E-S-G-A-E-L-E-K in region 303-370, 3 X 15 AA repeats of [DN]-[FS]-P-[STV]-R-S-H-[DE]-[FL]-D-[LM]-K-[NT]-E-[ST] in region 317-412, 5 X 5 AA repeats of [FV]-[ADT]-[EST]-[KM]-L in region 510-600, 2 X 23 AA repeats in region 648-696, 5 Asp and 4 Gly as compositional bias in position 63-67 and 638-341 respectively. DnaSP analysis for 5 *rd29A* from different species shown that for this gene singleton variable sites (two variants) is 135, Number of haplotypes is 5, Variance

Table 1: *Arabidopsis thaliana* promoter region motifs

Motifs	Position	Matrix score	Sequence	Function
3-AF1 binding site	532	10	AAGAGATATTT	Light responsive element
5UTR Py-rich stretch	118	9	TTTCTTCTCT	cis-acting element conferring high transcription levels
AE-box	1203	8	AGAAACTT	Part of a module for light response
ARE	380, 1358, 954, 765, 1135	6	TGGTTT	cis-acting regulatory element essential for the anaerobic induction
ATCT-motif	174, 1270	9	AATCTAATCT	Part of a conserved DNA module involved in light responsiveness
Box I	1364	7	TTTCAAA	Light responsive element
Box III	1391	9	CATTTACACT	Protein binding site
CAAT-box	83, 108, 177, 225, 240	4, 5	CAAT, CAAAT, CCAAT	Common cis-acting element in promoter and enhancer regions
CAT-box	880	6	GCCACT	cis-acting regulatory element related to meristem expression
TC-rich repeats	167, 1308, 1193	9	ATTTTCTTCA	cis-acting element involved in defense and stress responsiveness
TATA-box	1227 , 1388, 1407	6, 7	TATAAA, TATA, TATAA	Core promoter element around -30 of transcription start
MBS	24	6	TAACTG	MYB binding site involved in drought-inducibility
HSE	783, 1291	9	AAAAAATTTC, AGAAAATTCG	cis-acting element involved in heat stress responsiveness
GA-motif	880	8	AAAGATGA	Part of a light responsive element
WRKY71OS	1021	8	TGACY	transcriptional repressor of the DE gibberellin signaling pathway

UUU-F	24	(1.66)	UCU-S	23	(2.03)	UAU-S	5	(1.00)	UGU-C	0	(0.00)
UUC-F	5	(0.34)	UCC-S	8	(0.71)	UAC-Y	5	(1.00)	UGC-C	0	(0.00)
UUA-L	2	(0.38)	UCA-S	12	(1.06)	UAA-*	0	(0.00)	UGA-*	0	(0.00)
UUG-L	8	(1.50)	UCG-S	10	(0.88)	UAG-*	0	(0.00)	UGG-W	2	(1.00)
CUU-L	8	(1.50)	CCU-P	9	(0.73)	CAU-H	16	(1.68)	CGU-R	0	(0.00)
CUC-L	3	(0.56)	CCC-P	2	(0.16)	CAC-H	3	(0.32)	CGC-R	0	(0.00)
CUA-L	2	(0.38)	CCA-P	20	(1.63)	CAA-Q	12	(1.14)	CGA-R	1	(0.26)
CUG-L	9	(1.69)	CCG-P	18	(1.17)	CAG-Q	9	(0.86)	CGG-R	1	(0.26)
AUU-I	7	(1.75)	ACU-T	21	(1.62)	AAU-N	9	(1.06)	AGU-S	11	(0.97)
AUC-I	4	(1.00)	ACC-T	4	(0.31)	AAC-N	8	(0.94)	AGC-S	4	(0.35)
AUA-I	1	(0.25)	ACA-T	11	(0.85)	AAA-K	24	(0.87)	AGA-K	18	(1.70)
AUG-M	9	(1.00)	ACG-T	16	(0.23)	AAG-K	21	(1.13)	AGG-K	3	(0.78)
GUU-V	15	(1.20)	GCU-A	11	(1.47)	GAU-D	45	(1.50)	GGU-G	22	(1.38)
GUC-V	5	(0.40)	GCC-A	4	(0.53)	GAC-D	15	(0.50)	GGC-G	6	(1.38)
GUC-V	11	(0.88)	GCA-A	6	(0.80)	GAA-E	51	(0.94)	GGA-G	33	(2.06)
GUG-V	19	(1.52)	GCG-A	9	(1.20)	GAG-E	57	(1.06)	GGG-G	3	(0.19)

Fig. 1: Codon usage for rd29A gene in *Arabidopsis thaliana*, *Brassica oleracea*, *Thellungiella halophila*, *Craterostigma plantigineum* and *Glycine max*

of haplotype diversity is 0.01600, average number of nucleotide differences is 706.600, nucleotide diversity is 0.69275 and conserved regions are ATGGADYHRMVRNDDVHHYVHHSWNNHVDBV in position 1-31 and KVRAVVHMKMWKKKSBKSMRMA SSDRARDCCYRKDDRHGDHGASRMRKYWSHWKM KRAKWWWWYTHCDGMVRVHDDWVATRWBRWR RARRHWGMRHCWSHKVWRRVDRR in position 1271-1379. Also codon usage for all amino acid indicated in Fig. 1.

Protein alignment shows similarity and homology among species. Results indicated that *Arabidopsis thaliana rd29A* has high identity with corresponding genes in *Arabidopsis lyrata*, *Thellungiella halophila* and *Brassica oleracea* (Fig. 2). Also phylogenetic tree shown that this gene has more relative with *Arabidopsis lyrata*, *Thellungiella halophila* and *Cupressus sempervirens* (Fig. 3). Protein interaction defined by SMART program and investigated that *rd29A* as LTI78 (low-temperature-induced 78); cold regulated gene, the 5' area of cor78 has cis-acting regulatory elements that can impart cold-regulated gene expression (Fig. 4).

DISCUSSION

Environmental factors such as drought, salinity and cold stresses reduce plant yield and production. The *rd29A* is stress inducible gene which not expressed in normal and optimal conditions. Therefore in this research promoter and CDs regions of this gene analyzed. Results indicated that some stress inducible transcription factors and cis-acting regulatory elements such as HSE, MBS, TC-rich repeats, *etc* exist in *rd29A* promoter region, this phenomenon explain this gene expression under stress conditions. The promoters of *rd29A* and *rd29B*, which are 60.93% identical and the 1640 bp spacer region between *rd29B* and *rd29A* is thought to contain cis-acting elements responsible for the induction of *rd29A* by environmental stress and ABA. These observations suggest that the *rd29A* promoter region include at least two types of independent cis-acting elements, one which is responsive to changes in osmotic potential (dehydration-responsive element) and is independent of ABA and another which is, perhaps, involved in ABA-responsive (ABA-responsive element) gene expression (Yamaguchi-Shinozaki and Shinozaki, 1993a). By contrast, the *rd29B* promoter does not appear to contain cis-acting elements that are involved in the initial rapid induction caused by changes in osmotic potential.

Protein sequence analysis performed for finding motifs and functional domains. Also, Jia *et al*. (2012) reported that the middle region of *rd29A* is a sequence of 112 amino acids, composed of two tandemly repeated sequences. A second tandemly repeated sequence of 21 amino acids is located near the C-terminal. Both of these repeats are absent in *rd29B*. Two acidic regions and one basic region can be found in the N-terminal regions of *rd29A* and *rd29B*.

Our data analysis by STRING shown that *rd24A* has predicted protein interactions with KIN2; Encodes a gene that can be overexpressed by cold and abscisic acid and may be involved in temperature acclimation and salt

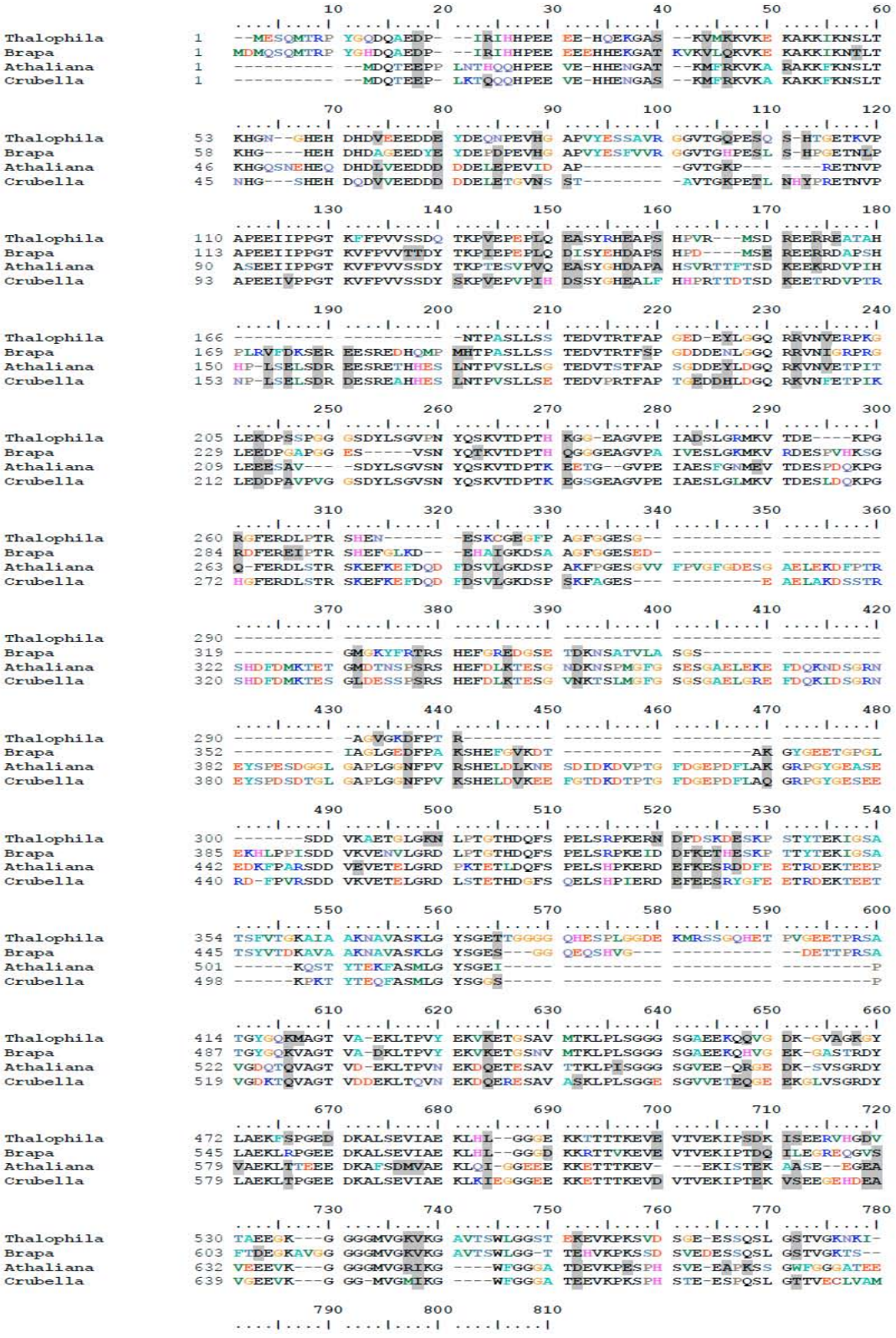

Fig. 2: Multiple alignment of rd29A protein among *Arabidopsis thaliana* and other species

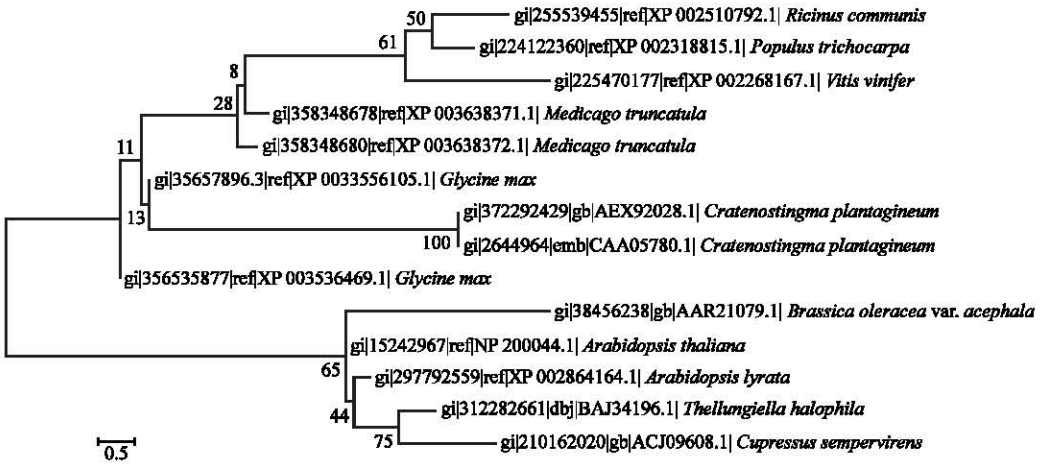

Fig. 3: Phylogenetic tree of *rd29A* genes from various species

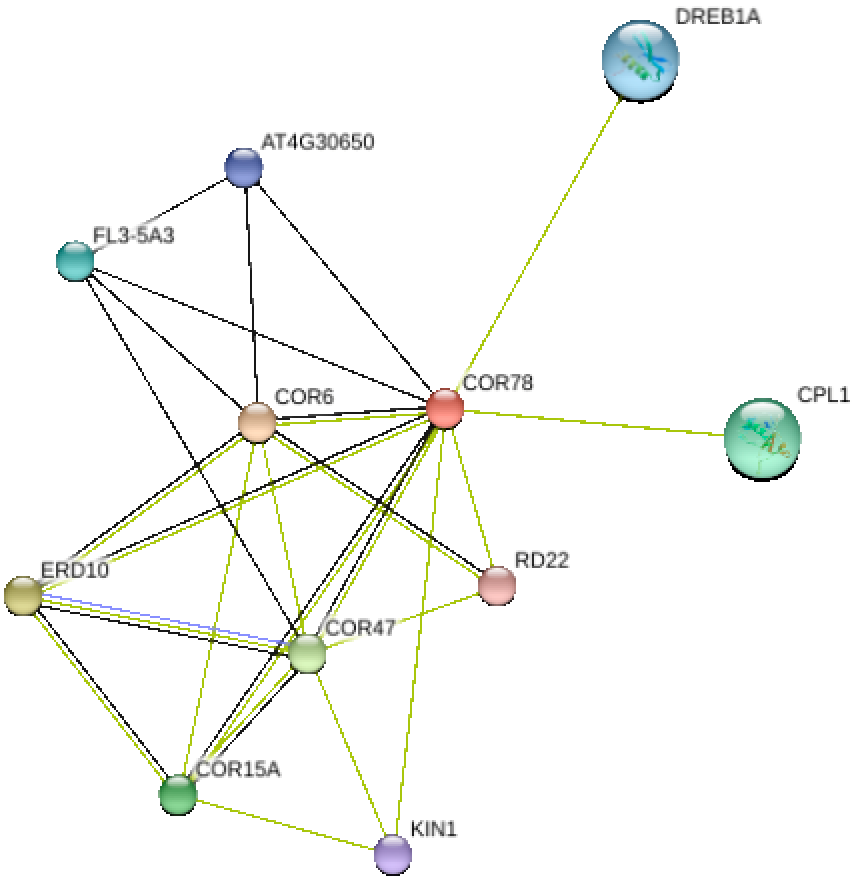

Fig. 4: The *rd29A* (COR78) predicted protein interactions

tolerance, ERD10 (Early Responsive to Dehydration 10); actin binding; Encodes a gene induced by cold, freezing and dehydration, COR47 (Cold-Regulated 47); Belongs to the dehydrin protein family, amino acid homology with LEA II proteins (Late Embryogenesis Abundant proteins), cold regulated and responds gene to osmotic stress, abscisic acid, drought, etc. (Bihmidine *et al*., 2013; Narusaka *et al*., 2003).

CONCLUSION

Tolerant to environmental stress is important for preserving food production in 21th century. Finding and transformating genes to improve adaptability of major crops such as wheat and rice are critical to adverse tensions environments. Experiments demonstrated that *rd29* genes expression increased in stress conditions. We analyzed *rd29A* gene and promoter and found some elementary regions and transcription factors related to stress condition. Generally, our results suggest that this gene is a good candidate for transformation into sensitive plants against different abiotic stress such as drought and temperature stresses.

REFERENCES

Bihmidine, S., J. Lin, J.M. Stone, T. Awada, J.E. Specht and T.E. Clemente, 2013. Activity of the *Arabidopsis* RD29A and RD29B promoter elements in soybean under water stress. Planta, 237: 55-64.

Jia, H., S. Zhang, M. Ruan, Y. Wang and C. Wang, 2012. Analysis and application of *RD29* genes in abiotic stress response. Acta Physiol. Plant, 34: 1239-1250.

Miller, F.P., A.F. Vandome and J. McBrewster, 2010. Arabidopsis Thaliana: Flowering Plant, Europe, Asia, Africa, Model Organism, Genome, Molecular Biology, Flower, Phototropism, Botany, Non-Mendelian Inheritance, Nottingham Arabidopsis Stock Centre. Alphascript Publishing, USA.

Modarresi, M., G.A. Nematzadeh, F. Moradian and S.M. Alavi, 2012. Identification and cloning of the Cu/Zn superoxide dismutase gene from halophyte plant *Aeluropus littoralis*. Russ J. Genet., 48: 118-122.

Modarresi, M., G. Nematzadeh and M. Zarein, 2013. Glyceraldehyde-3-phosphate dehydrogenase gene from halophyte *Aeluropus lagopoides*: Identification and characterization. J. Crop Improv., 27: 281-290.

Msanne, J., J. Lin, J. Stone and T. Awada, 2011. Characterization of abiotic stress-responsive *Arabidopsis thaliana* RD29A and RD29B genes and evaluation of transgenes. Planta, 234: 97-107.

Narusaka, Y., K. Nakashima, Z.K. Shinwari, Y. Sakuma and T. Furihata *et al*., 2003. Interaction between two cis-acting elements, ABRE and DRE, in ABA-dependent expression of Arabidopsis rd29A gene in response to dehydration and high-salinity stresses. Plant J., 34: 137-148.

Yamaguchi-Shinozaki, K. and K. Shinozaki, 1993a. Arabidopsis DNA encoding two desiccation-responsive rd29 genes. Plant Physiol., 101: 1119-1120.

Yamaguchi-Shinozaki, K. and K. Shinozaki, 1993b. Charactrization of the expression of a desiccation-responsive rd29 gene of Arabidopsis thaliana and analysis of its promoter in transgenic plants. Mol. Gen. Genet., 236: 331-340.

Yamaguchi-Shinozaki, K. and K. Shinozaki, 1994. A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature or high-salt stress. Plant Cell, 6: 251-264.