

***In silico* Identification and Analysis of Stress Responsive Elements Present on Promoter Region of *Acacia nilotica* ALDH7 gene**

Mahadev Asaram Jadhav¹, Kiran Ramesh Kharat¹ and Arun Sidram Kharat²

¹Department of Biotechnology and Bioinformatics, Deogiri College, Aurangabad, India

²School of Life Sciences, Jawaharlal Nehru University, New Delhi, India

*Corresponding author: majadhav22g@gmail.com

Abstract

The promoter region of *Acacia nilotica* ALDH7 gene consists of 1233 bp and it belongs to linear DNA molecular class. *A. nilotica* consist a universal feature to response moisture stress and nutrient limitation. In *A. nilotica* ALDH7 gene promoter sequence, the nucleotide motifs were involved in respond to various stressful conditions, but which regulatory elements in ALDH7 gene respond positively under stressful conditions unknown to till date. *In silico* approaches were used to estimate the nucleotide motifs and its response to certain stressful condition and transcription activation located at ALDH7 gene promoter region. The result reveals that, ALDH7 gene promoter region includes six unique signature motifs such as 'ACGTG' is a high salinity and induction of dehydration stress-responsive putative *cis*-regulatory motif, the 'TGAC' is the biotic and abiotic stress and transcription activation responsive motif, the 'CACGTG' is the defense responsive motif, the 'AAAG' is Adenosine rich motif involved in carbon metabolism, the 'GGATA & TAACAA' motifs responsible for transcription activation. Thus, our analysis suggests that ALDH7 gene promoter region play an important role to healthy survival of *A. nilotica* under stressful conditions and provides medicinal and infrastructural products to fulfil the requirements of human being.

Keywords: Motif; *Acacia nilotica*; ALDH7; Biotic & abiotic stress; Promoter region; Transcriptional Activation

Acacia nilotica is a multipurpose tree mainly distributed in the warm arid and semiarid regions of the world (El Amin, 1973; Tybirk, 1991). It is also a polyploid complex, which plays an essential role in the rural economy through the supply of wood, by-products, stabilization and the fertilization of soils (Fagg & Stewart, 1994; Wickens, 1995). Taxonomy was studied in detail by Brenan (Brenan, 1983) who divided the species in nine subspecies among which 3 are native from India and 6 from Africa. The subspecies can be differentiated mainly on the shape, the size and degree of pubescence of pods and shape of the crown. Transcriptional control is of critical importance in mediating the responses of eukaryotic cells to external stimuli. The promoter of a gene (the

regulatory DNA sequence upstream of the transcribed region) is centrally important in determining if and when transcription will be initiated. The nucleotide sequence of the promoter specifies the recruitment of DNA-binding proteins, including the transcription factors that regulate gene expression. The short DNA sequence motifs that specify protein binding are therefore the essential functional components of the promoter. The level of interest in the mechanisms of transcriptional regulation has led to a number of advances in the computational analysis of regulatory DNA sequence (Hudson & Quail, 2003).

In plants, a large number of transcription factors are known to control the expression of target genes in

various signal transduction cascades (Venter & Botha, 2004). Transcription-factor binding sites (TFBs or *cis*-regulatory elements), on the other hand, determine the specific timing and location of transcriptional activity. They are primarily located in the long non-coding sequence upstream of a gene (Chaboute, Clement & Phillips, 2002). These regulatory motifs organized into distinct *cis* regulatory modules are required for a specific expression pattern (Babu *et al.* 2004). Thus, the identification of regulatory motifs and their organization is an important step to improve understanding of gene expression and regulation.

Plants, being sessile, have evolved specific acclimation and adaptation mechanisms to respond to and survive short- and long-term drought stresses. Analysis of these protective mechanisms will contribute to our knowledge of tolerance and resistance to stress. The complex responses to environmental stress, from perception to transcriptional and physiological changes, need to be considered at a global systems biology level to study the multiple interactive components in this biological process (Krishnan & Pereira, 2008).

In recent years, much molecular information has been generated on the response of plants to environmental stresses. Plants respond to environmental stresses such as drought by the induction of both regulatory and functional sets of genes (Ingram & Bartels, 1996; Ramanjulu & Bartels, 2002; Bartels & Sunkar, 2005). Very little is known about the early events in the perception of stress signals (Urao *et al.* 1999; Ueguchi *et al.* 2001; Wohlbach, 2008). The identification of the stress-responsive *cis*-elements ABA-responsive element (ABRE) and dehydration-responsive element (Yamaguchi-Shinozaki & Shinozaki K, 1994). Downstream of the early signal perception events, signaling genes and molecules acting as secondary messengers have been identified, revealing the role of Ca⁺ and reactive oxygen species (ROS) as secondary messengers. These regulatory mechanisms induce downstream functional genes, which are needed to establish new cellular homeostasis that leads to

drought tolerance and/or resistance (Yancey, 2001). Water stress significantly reduced shoot growth in *Acacia nilotica* and *Leucaena leucocephala* (Michelsen and Rosendahl, 1990).

One way to identify the conditions upon which a *cis* sequence may confer gene expression is to analyse if genes harbouring this sequence in their promoter show a specific expression profile under certain environmental conditions (Kilian *et al.* 2007). The candidate *cis*-regulatory elements in the promoter regions can be identified by searching against the known elements in the databases. In addition, novel *cis*-regulatory elements could also be detected with no prior knowledge of transcription factor binding sites using the over-representative sequence of the promoters of co-expressed genes (Helden, 2003). Many databases have been developed for the prediction of *cis*-regulatory sites in a noncoding DNA sequence PLACE (Higo *et al.* 1999). Therefore, the use of *in silico* promoter analysis through bioinformatics as a tool has become more attractive and feasible when combined with experimental gene expression analysis.

MATERIALS AND METHODS

Retrieval of promoter sequence data

The nucleotide sequence of ALDH7 gene promoter in *Acacia nilotica* was retrieved from Genbank database. The data consists only sequence of promoter region in ALDH7 gene. GenBank is a comprehensive database that contains publicly available nucleotide sequences for almost 260000 formally described species (Benson, *et al.* 2013).

Searching of stress responsive element

The ALDH7 gene promoter sequence of *Acacia nilotica* was predicted using PLACE. The *cis* regulatory elements were listed and the occurrence number on each promoter was recorded. PLACE is a database of motifs found in plant *cis*-acting regulatory DNA elements. It covers vascular plants only (Higo *et al.* 1998).

RESULTS AND DISCUSSION

The ALDH7 gene promoter sequence comprises 1233 bp in single strand. Promoter region of ALDH7 gene in *Acacia nilotica* contains many *cis*-regulatory elements including stress related *cis*-acting elements. The putative responses of these important *cis*-regulatory elements which are involved in stresses, transcription activation, defense and carbon metabolism are recorded in (Table 1).

Plants, like other living organisms are exposed to various biotic and abiotic stresses and they have to cope with such situations. The expression of various proteins involved in stress responses is regulated through the interaction of transcription factors with upstream regions of genes. A number of stress related regulatory elements have been identified which are specifically involved in expression of their cognate genes. AAAG is a core site required for binding of Dof proteins, which are DNA binding proteins

Table 1: The nucleotide elements (motifs) and its putative response identified in ALDH7 gene promoter sequence in *Acacia nilotica*

Sl. No.	Nucleotide elements (motifs)/ Consensus sequence	Length in BP	Putative Response (function)
1	ACGTG	05	The induction of dehydration & high salinity stress
2	TGAC	04	The biotic and abiotic stress and transcription activation
3	CACGTG	06	Defence against pathogenic attack
4	AAAG	04	Involvement in carbon metabolism
5	GGATA	05	Transcription activation
6	TAACAA	06	Transcription activation

Table 2: The nucleotide elements (motifs) and its occurrence in various genes in different plant species

Sl. No	Nucleotide motifs Consensus sequence	Name of the genes where motif located and its source	Function or Mechanism of motif	Type of stress
1	ACGTG	ERD1, ACLPA, RD29A, RD29B (<i>A.thaliana</i>)	This motif necessary for Luciferase activity, seed germination and vegetative growth in dehydrated plants	Abiotic
2	TGAC	AMY32B (<i>O. sativa</i>) ERF3 (<i>N. tabacum</i>)	This motif found in DNA binding domain and it leads to highly diverged functions. Such as transcriptional activation & pathogenic interaction	Biotic and Abiotic
3	CACGTG	GBF4 (<i>A.thaliana</i>) CrGBF1, CrGBF2 (<i>C. roseus</i>) Pti4 (<i>L. esculentum</i>)	This motif act as a transcriptional repressor and it directly interact with G-box and it regulates defense related gene expression via GCC box which shows reduced disease symptoms.	Biotic and Abiotic
4	AAAG	DOF 1,2 (<i>Zea mays</i>)	Dof proteins may play regulatory roles in multiple gene expressions associated with the plant-specific pathway for carbon metabolism	Abiotic
5	GGATA	MYB (<i>S. tuberosum</i>) CAB22 (<i>P.hybrida</i>)	The repeats of this motif leads to increase the transcription level of genes occurred in etiolated seedlings and dark adapted plants in response to light.	Abiotic
6	TAACAA	GA 1,2,3 (<i>A.thaliana</i>)	Seed germination just before radicle emergence (GA hormone mediated)	Abiotic

involved in expression of photosynthetic genes, seed storage protein genes and genes responsive to various stresses (Yanagisawa & Schmidt, 1999; Kisu *et al.* 1998; Mena *et al.* 1998). The W box [(T) TGAC(C/T)] is the binding site for members of the WRKY family of transcription factors (Table 2).

There is increasing evidence that W boxes are a major class of *cis*-acting elements responsible for the pathogen inducibility of many plant genes (Wang *et al.* 1998) and, are also involved in senescence and trichome development (Eulgem *et al.* 1999). W box was also found to be involved as an enhancer in the induction of salicylic acid in response to wounding and bacterial infection in *Brassica oleracea*. The TGAC element was first reported in rice (Eulgem *et al.* 2000) and found to be involved in gibberellin signalling pathway. The TGAC motif is found as an invariant core sequence in W box which serves as a binding site for a rice transcription factor, WRKY71. Seven copies of TGAC are present in the promoter region of OsRGLP1 (Table 2).

The TTGAC/ TTGACC element is also part of W box and reportedly involved in pathogen inducibility in Parsley (Rushton, 1996). Two elements, CATGTG and CACATG reported to be involved in water stress (Simpson, 2003) response were also found on the promoter region with two copies of each, indicating that this promoter might be involved in water stress responses and all these motifs also present in various genes in different plant species (Table 2).

CONCLUSION

This analysis concluded that the regulatory elements seem to be associated with diverse functions and it plays an important role to healthy survival of *Acacia nilotica* under biotic and abiotic stress. A regulatory motif becomes major aid to develop stress tolerant plant varieties using genetic engineering.

Author's Contributions

The bioinformatics analysis & interpretation was the contribution of the first author. While the second & third author provided major support and expertise during the preparation of manuscript.

REFERENCES

- Babu, M.M., Luscombe, N.M., Aravind, L., Gerstein, M. and Teichmann, S.A. 2004. Structure and evolution of transcriptional regulatory network. *Curr. Opin. Struct. Biol.*, **14**: 283-291.
- Bartels, D. and Sunkar, R. 2005. Drought and salt tolerance in plants. *Crit. Rev. Plant Sci.*, **24**: 23-58.
- Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., Sayers Brenan, J.P.M. 1983. Manual on taxonomy of Acacia species. Present taxonomy of four species of Acacia (*A. albida*, *A. senegal*, *A. nilotica*, *A.tortilis*). FAO. Rome, 47.
- Chaboute, M.E., Clement, B. and Phillips, G. 2002. S phase and meristem-specific expression of tobacco RNR1 b gene is mediated by an E2F element located in the 5' leader sequence. *J. Biol. Chem.*, **277**: 17845-17851.
- El Amin, H.M. 1973. Sudan Acacias Bulletin, Forest Research Institute, Khartoum n° 1. 64.
- Eulgem, T., Rushton, P.J., Robatzek, S. and Somssich, I.E. 2000. The WRKY superfamily of plant transcription factors. *Trends Plant Sci.*, **5**: 199-206.
- EW, 2013. Genbank. Nucleic Acids Res 41(Database issue): D36-D42.
- Fagg, C.W. & Stewart, J.L. 1994. The value of Acacia and Prosopis in arid and semi-arid environment. *Journal of Arid Environments*, **27**(1): 3-25.
- Helden, J.V. 2003. Regulatory sequence analysis tools. *Nucl. Acids Res.*, **31**: 3593-3596.
- Higo, K., Ugawa, Y., Iwamoto, M. and Higo, H. 1998. PLACE: A database of plant *cis*-acting regulatory DNA elements. *Nucl. Acids Res.*, **26**(1): 358-359.
- Higo, K., Ugawa, Y., Iwamoto, M. and Korenaga, T. 1999. Plant *cis*-acting regulatory DNA elements (PLACE) database. *Nucl. Acids Res.*, **27**: 297-300.
- Hudson, M.E. and Quail, P.H. 2003. Identification of Promoter Motif involved in the Network of Phytochrome A-Regulated Gene Expression by Combined Analysis of Genomic Sequence and Microarray Data. *Plant Physiol.*, **133**(4): 1605-1616.
- Ingram, J. and Bartels, D. 1996. The molecular basis of dehydration tolerance in plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, **47**: 377-403.
- Kilian, J., Whitehead, D. and Horak, J. 2007. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J.*, **50**: 347-363.
- Kisu, Y., Ono, T., Shimofurutani, N., Suzuki, M. and Esaka, M. 1998. Characterization and expression of a new class of zinc finger protein that binds to silencer region of ascorbate oxidase gene. *Plant Cell Physiol.*, **39**: 1054-1064.

- Krishnan, A. and Pereira, A. 2008. Integrative approaches for mining transcriptional regulatory programs in Arabidopsis. *Brief Funct Genomics Proteomics*, **7**: 264–274.
- Mena, M., Vicente-Carbajisa, J., Schmidt, R.J. and Carbonero, P. 1998. An endosperm specific D of protein from barley, highly conserved in wheat, binds to and activates transcription from the prolamin box of a native b-hordein promoter in barley endosperm. *Plant J.*, **16**: 53-62.
- Michelsen, A. and Rosendahl, S. 1990. The effect of V.A. Mycorrhizal fungi, phosphorus and drought stress on growth of *Acacia nilotica* and *Leucaena leucocephala* seedlings. *Plant and Soil*, **124**: 7–13.
- Ramanjulu, S. and Bartels, D. 2002. Drought- and desiccation-induced modulation of gene expression in plants. *Plant Cell Environ*, **25**: 141–151.
- Rushton, P.J., Torres, J.T., Parniske, M., Wernert, P., Hahlbrock, K. and Somssich, I.E. 1996. Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *EMBO J.*, **15**: 5690-5700.
- Simpson, S.D., Nakashima, K., Narusaka, Y., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K.R.A. 2003. Two different novel *cis*-acting elements of *erd1*, a *clpA* Homologous Arabidopsis gene function in induction by dehydration stress and dark induced senescence. *Plant J.*, **33**: 259-270.
- Tybirk, K. 1991. Régénération des légumineuses ligneuses du Sahel. AAU Reports 28, Botanical Institute Aarhus University, Denmark 85.
- Ueguchi, C., Koizumi, H., Suzuki, T. and Mizuno, T. 2001. Novel family of sensor histidine kinase genes in Arabidopsis thaliana. *Plant Cell Physiol.*, **42**: 231–235.
- Urao, T., Yakubov, B., Satoh, R., Yamaguchi-Shinozaki, K., Seki, M., Hirayama, T. and Shinozaki, K. 1999. A transmembrane hybrid-type histidine kinase in Arabidopsis functions as an osmosensor. *Plant Cell*, **11**: 1743–1754.
- Venter, M. and Botha, F.C. 2004. Promoter analysis and transcription profiling: Integration of genetic data enhances understanding of gene expression. *Physiol. Plant*, **120**: 74-83.
- Wang, Z., Yang, P., Fan, B. and Chen, Z. 1998. An oligo selection procedure for identification of sequence-specific DNA-binding activities associated with plant defense. *Plant J.*, **16**: 515-522.
- Eulgem, T., Rushton, P.J., Schmelzer, E., Hahlbrock, K. and Somssich, I.E. 1999. Early nuclear events in plant defense signaling: rapid gene activation by WRKY transcription factors. *EMBO J.*, **18**: 4689-4699.
- Wickens, G.E. 1995. Role of Acacia species in the rural economy of dry Africa and the Near East. FAO Conservation Guide, N° 27, FAO Rome.
- Wohlbach, D.J., Quirino, B.F. and Sussman, M.R. 2008. Analysis of the Arabidopsis histidine kinase ATHK1 reveals a connection between vegetative osmotic stress sensing and seed maturation. *Plant Cell*, **20**: 1101–1117.
- Yamaguchi-Shinozaki, K. and Shinozaki, K. 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.
- Yanagisawa, S. and Schmidt, R.J. 1999. Diversity and similarity among recognition sequences of D of transcription factors. *Plant J.*, **17**: 209-214.
- Yancey, P. 2001. Water stress, osmolytes and proteins. *Am. Zool.*, **41**: 699–709.

