The Quarterly Scientific Journal of Applied Biology
Vol 35, No. 3, Autumn, 2022, P: 2-5
Journal homepage: https://jab.alzahra.ac.ir
10.22051/JAB.2021.34043.1392

Evolutionary relationships of *ATAXIN-2* protein domain structure: A possible insight into the involvement of this protein in localized translation in rice embryogenesis process

Sh. heidari[1], P. heidari[2], B. Heidari[3*]

1, 2. Dept. of Plant Breeding, Science and Research Branch, Islamic Azad University, Tehran, Iran
3. Dept. of Computer Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran
(*Corresponding author: b_heidari@iauksh.ac.ir)

Received: 2021.09.19 **Accepted: 2022.05.29**

Introduction

With the development of high-throughput sequencing technologies, new sequence data are collected at a high speed, while annotating their function is a major challenge (Wong *et al*., 2019). Laboratory methods to determine protein function are a difficult step and a major limiting process for determining function. Traditional bioinformatics approaches, such as protein function annotation following database searches to find protein sequence homologues, are unable to assign appropriate functions to uncharacterized proteins (Engelhardt *et al*., 2009). A protein function determination method includes functional evidence from all members of a family, which reflects the evolutionary path of a set of homologous proteins using a phylogenetic tree and makes strong predictions for unknown proteins. Due to the modular restructuring of domains during protein evolution known as "domain shuffling", most proteins are composed of multiple domains, which results in higher pairwise homology relationships between single domain sequences than complete protein sequences. Taking advantage of this fact can increase the sensitivity of protein function prediction approaches (Chothia *et al*., 2003). *ATAXIN-2* consists of two globular domains named Lsm and LsmAD and also harbors a C-terminal motif (PAM2) at the carboxy-terminus. The domain structure of the *ATAXIN-2* protein provides information about its molecular function in RNA metabolism. In their own study, Heidari (2010) reported *ATAXIN-2* as a protein-coding gene with an unknown function and a significant increase in expression in rice embryogenesis. Many studies have reported increased expression of this protein in embryonic tissues of different organisms (Kiehl *et al*., 2006; Kiehl *et al*., 2000). So far, the exact cellular function of *ATXN2* in plants remains

unknown. Therefore, in this research, the evolutionary relationships of *ATXN2* domains in different plant species were used to predict the function of this unknown protein.

Methods and Materials

In order to investigate the conserved domains in the *ATAXIN-2* family in plants, all members of the *ATAXIN-2* family that are registered in PlantGDB were used as query sequences for searching in Pfam (Finn *et al*., 2016). A Pfam scan was performed using hmmscan, from the HMMer software suite (Eddy, 2011), using the gathering threshold. For the phylogenetic analysis in order to investigate the relative divergence in the large *ATAXIN-2* family, the Phylip programme (Retief, 2000) was used. Multiple alignments were generated by MAFFT using the L-INS-i method (Katoh & Standley, 2013). The phylogenetic tree was constructed using the Phylip collection using NEIGHBOR, FITCH, and PROML programmes with 100 bootstrap replicates. In order to conserve functional residues, the complete sequence of all proteins in the *ATAXIN-2* superfamily was used to perform MEME in order to search for the 3 superior motifs with lengths between 20 and 60 residues. Out of the top 5 identified motifs in the initial searches, only three samples had a MEME E-value $<10^{-100}$, and of these 3, motifs identified by MEME were used to run MAST and to locate the motif (E-value $<10^{-5}$) to identify other proteins containing these motifs. Finally, the function of *ATAXIN-2* protein-interacting partners was predicted using the STRING website (von Mering *et al*., 2003).

Results & Discussion

Pfam search results with members of the *ATAXIN-2* protein family were used to study the domain structure. The *ATAXIN-2* protein family in plants has various combinations of recognizable Pfam domains. Collections of 164 sequences containing at least one of the 3 main domains, namely the SM domain, LsmAD domain, and C-terminal, were recovered from 45 plant species. Putative orthologues were identified in most species, indicating that *ATAXIN-2* are important genes in plant genomes. A phylogenetic tree based on 3 conserved domains showed that all *ATAXIN-2* sequences in plants are divided into two main categories. One clade included class I and the other clade included class II. In class I, most of the sequences had two main conserved motifs, i.e., the SM domain and the C-terminal motif. In class II, most of the sequences had all three conserved motifs. Most of the studied species had both class I and class II types of the *ATAXIN-2* protein. The presence of both class I and II types of *ATAXIN-2* protein in most of the studied species indicates the strong conservation of both types of sequences and probably shows the importance of the function of this protein family in plants. The results of predicting the function of unknown protein in rice plant along with the prediction of the function of their interacting partners by the STRING site determined the molecular function of *ATAXIN-2* protein as RNA-binding and their biological function as regulation of cytoplasmic mRNA. Lastres-Becker *et al* (2008) state the main role of the *ATAXIN-2* protein in the broad modulation of local translation of mRNA. This role of *ATAXIN-2* probably refers to an alternative mechanism for protein localization in which protein-encoding mRNA is targeted and subsequently translated in situ. The most common mechanism for protein localization involves direct targeting of the protein itself through specific sequences such as nuclear or mitochondrial localization sequences. Many cellular proteins are located in specific subcellular locations. Spatial localization provides the possibility of functional division. Significant advances in RNA detection methods have led to the identification of an increasing

number of localised mRNAs. Until the beginning of the last decade, the described target mRNA sets were approximately 100, and the mRNA localization process was limited to specific cells. However, recent genome-wide analyses have dramatically changed this view, showing that subcellular targeting of mRNAs is a common mechanism used by polar cells to establish specific functional compartments (Palacios & Johnston, 2001). Also, hundreds to thousands of mRNAs have been detected in various cell compartments such as the mitotic apparatus, pseudopodia, and highly specialised cells such as neurons, dendrites, or axons (Medioni *et al*., 2012). Intracellular mRNA targeting has expanded with the identification of localised mRNAs in a wide range of organisms other than animals, including yeast, bacteria, fungi, and plants. However, not much is known about the role of RNA localization mechanisms in specific plant cell functions. Heidari (2010) in the study of gene expression and functional genomics of different morphogenesis pathways using the analysis of EST sources, reported a significant increase in the expression of the EST sequence attributed to *ATAXIN-2* in rice embryogenesis. The establishment of the apex-basal axis is an important event in plant embryogenesis, which is evident from the first stages of the embryo sac and zygote. There is now more evidence that specific types of signalling play an important role in coordinating gene expression programmes to establish polarity. Thus, mRNAs are more than simple "messengers" that convey genetic information from DNA to the protein synthesis machinery, where functionally relevant mRNAs can be simultaneously translated according to biological needs, providing efficient signalling to coordinately control gene expression. It has been found that in the embryo, through the local expression of key genes, two distinct functional meristems are formed in each pole. The algae Fucus provides some experimental features in studying the early events of zygote polarization. It has been observed that along with axis formation, localization or redistribution of plasma membrane components, including ion channels, redistribution of calcium to basal ends, localization of F-actin in the rhizoderm, the asymmetric distribution of RNA molecules in the zygote and a polar secretion of Golgi-derived cell wall components towards the "basal" region occurs (Shaw & Quatrano, 1996). This RNA-based mechanism involves the coordination of multiple complex processes, including mRNA transfer, targeting, and translation, and enables precise stimulus-induced control over protein position, abundance, and function. Therefore, it can be argued that in plants, like other organisms, different sets of mRNA are targeted for different parts and lead cells to respond to different stimuli and reprogramming towards proliferation, differentiation, growth, apoptosis, etc. This new layer of intracellular patterning, which was initially thought to be unique to highly specialised cells, may be widely present in many cell types, highlighting the role of *ATAXIN-2* in this field in plants. This review shows the possible role of *ATAXIN-2* in localised translation as a distinct mode of gene expression control that programmes gene function with precision, efficiency, and spatio-temporal flexibility.

Conclusion

In this study, the function of *ATAXIN-2*, which plays a role in the embryogenesis of rice plant, was determined as RNA-binding protein, which plays a role in the regulation of cytoplasmic mRNA. The role of these proteins may be in targeting individual mRNA subsets to subcellular locations where they await a signal to continue protein synthesis according to intrinsic and extrinsic conditions. Local translation of mRNAs instead of proteins has several significant advantages for a bipolar embryo, including reduction of transfer costs, preventing gene expression before mRNA reaches the appropriate location because, in a number of organisms, inappropriate spatial expression disrupts

the embryonic pattern, changing the concentration of proteins in macromolecular complexes, and facilitating the simultaneous translation of different subunits. Therefore, the *ATAXIN-2* protein can be isolated and targeted for further laboratory studies.

Keywords: Local translation, Phylogenetics, Prediction of function, RNA-binding proteins

Acknowledgement*:* The authors of the article are grateful to the Islamic Azad University of Kermanshah for supporting the current research.

Declaration of conflict of interest: The authors declare that they have no conflicts of interest.