RAWATSAR P.G. COLLEGE
'Sanskriti Ka Badlta Swaroop Aur AI Ki Bhumika' (SBSAIB-2025)
DATE: 25 January 2025
International Advance Journal of Engineering, Science and Management (IAJESM)
Multidisciplinary, Multilingual, Indexed, Double-Blind, Open Access, Peer-Reviewed,
Refereed-International Journal, Impact factor (SJIF) = 8.152

In silico study of osmotic stress responsive genomic elements with special reference to CKIN/SnRK genes and bZIP TFs family in Pigeon pea (Cajanuscajan L.)

Arti, Research Scholar, Department of Botany, Asian International University, Imphal West, Manipur

Introduction

Pigeon pea is a leguminous crop commonly grown in India, and is primarily cultivated in tropical regions. Although it is often overlooked, this crop has immense potential for improvement in both quantity and quality of production in India. Among all the legumes grown in the region, pigeon pea stands out for its unique combination of optimal nutritional profiles, high tolerance to environmental stress, high biomass productivity, and its ability to contribute essential nutrients and moisture to the soil. Pigeon pea is a rich source of starch, protein, calcium, manganese, crude fiber, fat, trace elements, and minerals. Moreover, apart from its high nutritional value, pigeon pea is also used in traditional folk medicine in India, China, Philippines, and other countries (Sharma et. al 2011).

Proper nutrition is a critical foundational requirement, as it has a significant impact on health, work performance, and cognitive development. However, in many developing nations, there is a growing prevalence of hunger and malnutrition (FAO, 1980). This can be attributed to factors such as rapid population growth, scarcity of arable land, and elevated food costs, as reported by the Food and Agriculture Organization (Pelletier et al. 1995). Legumes are a valuable source of nutrition, containing high levels of protein, energy, and essential vitamins and minerals. Researchers Vadiveli and Janardhanan noted the nutritional significance of legumes in 2005. Pigeon pea, a type of legume belonging to the Leguminosae family, is particularly prevalent in India, where it is widely grown and consumed. Pigeon pea is also known by various other names such as red gram, arhar, and tur dal, as highlighted by (Ghadge et al. 2008).

For a long time, pigeon pea was believed to be one of only two species in the Cajanus DC genus. However, further research has shown that this genus is now considered to be part of a larger group of related genera, including Atylosia, Endomallus, Rhynchosia, and Dunbaria. As a result, the Cajanus genus has expanded to include a total of 32 distinct species.

Pigeon pea is a leguminous shrub that can grow up to 5 meters in height. It likely originated in South Asia and first appeared in West Africa around 2000 BC, which is now recognized as a second major center of origin. Pigeon pea was brought to the West Indies through the slave trade, where it gained the nickname "pigeon pea" due to its use as bird feed in 1692, according to (Van der Maesen's 1985). The plant's leaves are arranged spirally on the stem and consist of three leaflets. Flowers are typically yellow and grow in terminal or axillary racemes, measuring 2-3 cm in length. Although pigeon pea pods are usually green, they can be hairy, streaked, or even colored dark purple, and contain 2-9 seeds per pod, which can vary widely in color and weight between 4-25g/100 seeds (Sheldrake, 1984). While the most common use for pigeon pea is in the form of dry split seeds (dhal), the pods and seeds are also harvested and consumed as a green vegetable in many countries.

Pigeonpea is a crucial legume food crop that has been cultivated for more than 3,500 years in India. It is grown on approximately 5 million hectares and is the sixth most significant legume food crop globally (Mula et.al 2010). This crop is a vital protein source for more than a billion people in the developing world and supports the livelihoods of millions of resource-poor farmers across Asia, Africa, South America, Central America, and the Caribbean (Royes et.al Longmans, London and New York, 1976). The challenge of protein deficiency in the developing world, where protein is often available at levels less than one-third of the minimum dietary requirements, is likely to worsen due to the increasing human population and stagnation in crop yields. Legumes, such as pigeonpea, provide a highly nutritious and balanced source of calories and protein that is not provided by commonly grown cereals in semi-arid regions, making them crucial for food security(Varshney et. al 2012). Major efforts are underway worldwide to increase Pigeon pea production by genetic analysis of the key traits, and then utilizing the genomic resources, which can accelerate progress. Increased population coupled with consumption preferences has resulted in a substantial increase in the demand for Pigeon pea. However, there is an urgent need to increase the yield and productivity to feed the burgeoning world population along with increased tolerance against various biotic and abiotic stresses limiting wheat production by significantly decreasing crop yield (Joshi et al. 2017). Osmotic stress is however a major challenge for wheat production as it causes devastating effects on crop yield. Therefore, finding ways to improve crop tolerance to osmotic stresses is essential to improve Pigeon pea productivity to achieve food security.

Definition of problem

Designing of crops with improved tolerance against ionic and hyperosmotic constraints is need of the hour (Gupta et al. 2015; Joshi et al. 2018). However, it is often difficult, due to highly variable environmental conditions. Plants have evolved different acclimation and adaptation mechanisms for multiple stress tolerance (Shoeva et al. 2017; Gupta et al. 2017). Once the stress is recognized, response regulator genes re-establish organismal and cellular homeostasis or decline periodic shock effects (Mickelbart et al. 2015). Timely and precise activation of these defense systems before the onset of damage is of prime importance for the endurance of plants under stressful conditions (Gupta et al. 2015). Therefore, precise responses being transmitted by a network of signaling molecules are vital. Given the complex genetic architecture of wheat necessitates the consideration of natural variation in genetic architecture as an integral part to identify key regulatory players (Akpinar et al. 2018). However, the use of these natural variations for crop improvement is still lagging due to lack of knowledge of their related biological mechanisms and associated genes.

The availability of complete genomic sequences of Pigeon pea in public domain has provided a unique opportunity for study of transcriptional regulatory networks controlled by sequence-specific DNA-binding transcription factors (TFs) which bind to DNA and either activate or repress gene transcription. The specific interactions between TFs and their binding sites, i.e. the *cis*-regulatory sequences, play a central role in the regulation of different biological processes such as development, growth, cell division, and responses to environmental stimuli. Identification, characterization, and annotation of TF repertoires from wheat will provide an insight on TF organization and biological functions of the TFs osmotic tolerance as well as their evolution.

The basic leucine zipper (bZIP) represents one of the largest as well as most diverse transcription factor (TFs) families. Transcription factors of the basic leucine zipper (bZIP) family control important processes in all eukaryotes. Inplants, bZIPs are regulators of many central developmental and physiological processes including photomorphogenesis,leaf and seed formation, energy homeostasis, and abiotic and biotic stress responses.The CKIN/SnRK gene family plays an important role in energy sensing and stress-adaptive responses in plant systems.

Review of literature

Pigeon pea is grown in diverse agro-climatic zones and is routinely confronted by several stresses including water scarcity, hyper salinity, extreme temperatures, heavy metals, UV radiation, herbivores and pathogens, severely affecting the crop yield and quality. Primary reason behind this loss is the oxidative damage to biomolecules caused by increased accumulation of reactive oxygen species (ROS), resulting in severe cellular impairment

(Suzuki et al. 2012; Keunen et al. 2013; Komatsu et al. 2014; Khare et al. 2015). Plants have evolved complex physiological, cellular and molecular machineries for responses and adaptations to environmental cues including osmotic stress (Kumar et al. 2017). When plants encounter adverse environmental conditions such as osmotic stress, a variety of genes regulating different functions are either upregulated or downregulated to confer stress tolerance in plants (Bhattacharjee et al. 2016). Regulation of various processes at transcriptional level plays crucial roles in controlling myriad of plant biological processes and signaling transduction cascades, and interaction of transcription factors (TFs) with promoter binding region of target genes either repress or activate downstream genes (Shang et al. 2013). TFs play a key role in modulating stress-sensory pathways and precise response towards stress (Deinlein et al. 2014). Several families of stress responsive TFs, including ERF/AP2, bZIP, MYB, MYC, NAC and WRKY have been well characterized for their regulatory roles in eliciting stress-responses in plants (Chen et al. 2015; Joshi et al.2016; Wani et al. 2016). Since transcriptional control of the expression of stress-responsive genes is a critical part of the plant response to stress, TFs have received great attention from plant physiologists to decipher their roles and plant biotechnologists to target them for producing stress tolerant plants. Numerous transgenic plants with improved tolerance against abiotic stresses have been developed through genetic engineering and have been reported previously. In recent years, with the momentum the non-coding RNAs have gain, the importance of TFs have scaled new heights as they are identified to be major targets of non-coding RNAs for growth modulation under abiotic stress conditions. Often abiotic oxidative stress responses require altered expression of stress-responsive genes and heat shock proteins, which are tightly regulated by TF network including HSFs (Guo et al. 2016). Several abiotic stress responsive genes have been identified from plants due to the advent of molecular techniques including large-scale transcriptome analyses. These stress-responsive genes perform wide spectrum roles ranging from cell protection through the production of regulatory functional proteins (such as TFs) to regulate signal transduction and gene expression under stress conditions (Budak et al. 2015). TFs are essential regulatory proteins as they play critical roles in converting stress signals perceptions to stress-regulated gene expression via their interactions with cis-regulatory elements present in their promoter (Budak et al. 2015) and thus, ultimately activating a signaling cascade of a network of genes which defines the abiotic stress tolerance of plants (Akhtar et al. 2012; Glover-Cutter et al. 2014). Several TF genes of wheat have been implicated in response to different abiotic stresses (Budak et al. 2013). Delineating regulatory behavior of altered gene expression at post-transcription level via micro RNAs (miRNAs) is of utmost importance to unravel the tolerance mechanism (Alptekin et al. 2017). Apart from several studies showing the differential expression of miRNAs and their targets to cope up osmotic stress in Pigeon pea, the study is still underperformed due to its inadequate genome access.

During the response and adaptation to various abiotic stresses, transcription factors act as triggers of gene expression and play important regulatory roles (Xiang et al., 2008). The basic leucinezipper (bZIP) proteins make up one of the largest and most diverse transcription factor families in plants, and regulate various biological processes, including seed maturation and germination, photomorphogenesis, floral induction and development, and pathogen defense (Jakoby et al., 2002; Nijhawan et al., 2008). The bZIP proteins possess an eponymous bZIP domain, which is comprised of a basic region and a leucine zipper (Hurst, 1995).

The first described SnRK family member was the yeast Snf1, being well-known by its role in lipid accumulation(Kamisaka et al., 2007)and glucose repression, regulating carbon metabolism(Baena-González et al., 2008)SnRK family has been considered a potential target to improve plant performance under unfavorable conditions.(Coello et al. 2008). The

Chlorophyceae*Chlamydomonasreinhardtii*, thereafter *Chlamydomonas*, shares common ancestry with vascular plants. (Valledor et al. , 2014) Therefore, it is expected that many of its responses to limiting conditions would be similar (Valledor et al., 2013). Few studies suggest the involvement of SnRK family, named CKIN in *Chlamydomonas*, in stress response(Valledor et al., 2013).

Objectives

1. Genome-wide identification of osmotic stress-responsivebZIP TFs family in Pigeon pea genome.
2. Genome-wide identification of osmotic stress-responsive CKIN/SnRKgenes in Pigeon pea genome.
3. *In-silico*based functional and structural insights to thebZIP TFs family and CKIN/SnRK candidate genes involved in response to osmotic stress response.
4. Homology modeling of putativebZIP TFs and CKIN/SnRK osmotic stress responsive proteins.

Methodology

Data collection:

Amino acid sequences of reference genes and TFs will be obtained from different biological databases including NCBI, Uniprot,TAIR, Gramene RAP-DB etc.

Identification and chromosome localization of genes and TFs in wheat genome:

BLASTP algorithm will be carried out against latest Pigeon pea genome assemblies available on IWGSC-URGI (https://urgi.versailles.inra.fr/) and Ensembl Plants (http://plants.ensembl.org/index.html) databases.

Prediction of exon and Intron boundaries:

The Intron/Exon boundaries for individual genes and will be determined by aligning the CDS sequences to their corresponding genomic DNA sequences that will be utilized as input for graphical display at the Gene Structure Display Server (GSDS v2.0) of Peking University, China (http://gsds.cbi.pku.edu.cn/).

Functional domains, physicochemical properties and subcellularlocalization analysis:

Conserved domain database (CDD) of NCBI (www.ncbi.nlm.nih.gov) will be utilized to predict the functional domains of osmotic stress responsive genes and TFs. Physicochemical properties including molecular weight (MW), isoelectric point (pI), charge and average residue weight, will be identified by using ProtParam server of Expasy tools (https://web.expasy.org/protparam/)while subcellular localization will be predicted by using consensus algorithms of TargetP1.1 (http://www.cbs.dtu.dk/services/TargetP/)and PLANT-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/).

Phylogeny analysis:

Phylogeny analysis will be performed by multiple sequence alignment (MSA) using ClustalX ver. 2.1. MEGA 7.0 will utilize to obtain the phylogenetic tree using the neighbor-joining method.

Prediction of molecular markers and targeting miRNAs:

Molecular markers, simple sequence repeats (SSRs) will be predicted in genomic transcripts of identified wheat genes using BatchPrimer3v1.0 (http://probes.pw.usda.gov/batchprimer3/). Web-based psRNA Target server (https://plantgrn.noble.org/psRNATarget/email?addcdnalib=yes) will be used to predict the targeting miRNAs with default parameters.

Homology modeling:

Homology modeling method will be used to predict the 3D structures of identified proteins. The position specific iterated BLAST (PSI-BLAST) against Protein Data Bank (PDB) (http://www.rcsb.org/pdb/home/home.do) will be carried out to identify the homologous template structures of identified proteins, whereas 3D structure will be predicted using automated based Swiss-Model server (https://swissmodel.expasy.org/). Further, UCSF CHIMERA 1.10 (https://www.cgl.ucsf.edu/chimera/) will be used to obtain in different chemical coordinates of modeled proteins.

CONCLUSION

The identified genes and TFs that improve the osmotic stress tolerance crop plants and the candidate TF genes existing in pathways relating to osmotic stress signaling and regulation. Identifying osmotic stress tolerance genes will facilitate the breeding of crops with improved varieties. As such, systematic bioinformatics approaches to efficiently utilize such as the development of efficient crop varieties. This strategy could lead to high yields with improved quality of crop varieties.Through this study, generated genomic information on TF gene in Pigeon pea can be used in breeding for better osmotic stress tolerance crop varieties.

The bZIP family of transcription factors is a highly diverse and extensive group that plays a crucial role in controlling various biological processes in all eukaryotes. These TFs are particularly significant in plants, where they act as key regulators of essential developmental and physiological processes such as photomorphogenesis, seed and leaf formation, energy balance, and responses to both biotic and abiotic stresses. The CKIN/SnRK gene family is also of great importance in plants as it helps in detecting energy levels and enables adaptive responses to environmental stresses.

References:

1. Agriculture Organization. The double burden of malnutrition: case studies from six developing countries. Food & Agriculture Org.; 2006.
2. Alptekin, B., Langridge, P., Budak, H. Abiotic stress miRNomes in the Triticeae. FunctIntegrGenom,17(2–3),145–170.
3. Araujo J, Díaz-Alcántara CA, Velázquez E, Urbano B, González-Andrés F. Bradyrhizobiumyuanmingense related strains form nitrogen-fixing symbiosis with Cajanuscajan L. in Dominican Republic and are efficient biofertilizers to replace N fertilization. ScientiaHorticulturae. 2015 Aug 31;192:421-8.
4. Baena-González, E. Sheen, J. (2008). Convergent energy and stress signaling. Trends Plant Sci. ,13, 474–82 .
5. Coello, P., Hey, S. J. Halford, N. G. (2011). The sucrose non-fermenting-1-related (SnRK) family of protein kinases: potential for manipulation to improve stress tolerance and increase yield. J. Exp. Bot. 62, 883–93.
6. Duvick, J., Fu, A., Muppirala, U., Sabharwal, M., Wilkerson, M.D., Lawrence, C.J., Lushbough, C., Brendel, V.(2008).PlantGDB: a resource for comparative plant genomics. Nucleic Acids Res. 36: 959-965.
7. Fang, Y., You, J.,Xie, K., Xie, W. and Xiong, L. (2008). Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice, Mol. Genet. Genomics, 280, 547–63.
8. Fredslund, J. (2008). DATFAP: a database of primers and homology alignments for transcription factors from 13 plant species, BMC Genomics, 9, 140
9. FrongilloJr EA. Symposium: causes and etiology of stunting. J Nutr. 1999;129(2S Suppl):S529-30.
10. Gao, G., Zhong, Y., Guo, A., et al. (2006). DRTF: a database of rice transcription factors, Bioinformatics, 22, 1286–7.
11. Ghadge PN, SHEWALKAR SV, Wankhede DB. Effect of processing methods on qualities of instant whole legume: Pigeon pea (Cajanuscajan L.). Agricultural

Engineering International: CIGR Journal. 2008.

12. Guo, A.Y., Chen, X., Gao, G., et al. (2008).PlantTFDB: a comprehensive plant transcription factor database, Nucleic Acids Res., 36, D966–9.
13. Hurst, H. (1995). Transcription factors 1: bZIP proteins. Protein Profle 1, 101–168.
14. Hutchinson J, editor. Evolutionary studies in world crops: Diversity and change in the Indian subcontinent. CUP Archive; 1974 Apr 25.
15. Ichikawa, T., Nakazawa, M., Kawashima, M., et al., (2003). Sequence database of 1172 T-DNA insertion sites in *Arabidopsis* activation-tagging lines that showed phenotypes in T1 generation, Plant J., 36, 421–9.
16. Iida, K., Seki, M., Sakurai, T., et al. (2005). RARTF: database and tools for complete sets of Arabidopsis transcription factors, DNA Res., 12, 247–56.
17. Jakoby, M. et al. (2002).bZIP transcription factors in Arabidopsis. Trends Plant Sci. 7, 106–111.
18. Jin, J., Zhang, H., Kong, L., Gao, G., Luo, J. (2014). PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. Nucleic Acids Res ,42 , 1182-1187.
19. Johnson, C., Bowman, L., Adai, A.T., Vance, V., Sundaresan, V. (2007). CSRDB: a small RNA integrated database and browser resource for cereals. Nucleic Acids Res.,35,829-833.
20. Kumar, A., Kumar, A., Tyagi, P., Singh, K.(2018). Identification and Functional Characterization an Ortholog of OsENOD93-1 Gene in Wheat Using in-silico Approach. International Journal of Current Microbiology and Applied Sciences., 7, 3240-3250.
21. Kumari, S., Kumar, A., Tuteja, N.(2018). Genome-wide Comprehensive Analysis and Abiotic Stress Responsive Expression Profiles of MCM Gene Family in Wheat. International Journal of Current Microbiology and Applied Sciences. 7. 940-952.
22. Künne, C., Lange, M., Funke, T., Miehe, H., Thiel, T., Grosse, I., Scholz, U. (2005).CR-EST: a resource for crop ESTs. Nucleic Acids Res.,33, 619-621.
23. Kuromori, T., Takahashi, S., Kondou, Y., Shinozaki, K. and Matsui, M., (2009). Phenome analysis in plant species using loss-of-function and gain-of-function mutants, Plant Cell Physiol., 50, 1215–31.
24. Latham, M.C. Human nutrition in the developing world. FAO Food and Nutrition Series No. 29 (UN Food and Agriculture Organization, 1997)
25. Mochida, K., Yoshida, T., Sakurai, T., Yamaguchi-Shinozaki, K., Shinozaki, K., Tran, L.S. (2011). *In silico* analysis of transcription factor repertoires and prediction of stress-responsive transcription factors from six major gramineae plants. DNA Res. 321-332 .
26. Mula MG, Saxena KB. Lifting the level of awareness on pigeonpea-a global perspective. International Crops Research Institute for the Semi-Arid Tropics; 2010.
27. Nijhawan, A., Jain, M., Tyagi, A. K. Khurana, J. P. (2008). Genomic survey and gene expression analysis of the basic leucine zipper transcription factor family in rice. Plant Physiol. 146, 333–350.
28. Royes, W.V. Pigeonpea. in *Evolution of Crop Plants* (ed., Sommonds, N.W.). 154–156 (Longmans, London and New York, 1976).
29. Schroeder, J.I., Delhaize, E., Frommer, W.B., Guerinot, M.L., Harrison, M.J., Herrera-Estrella, L., Horie, T., Kochian, L.V., Munns, R., Nishizawa, N.K., Tsay, Y.F., Sanders, D.(2013). Using membrane transporters to improve crops for sustainable food production. Nature ,497, 60-66.
30. Sharma S, Agarwal N, Verma P. Pigeon pea (Cajanuscajan L.): a hidden treasure of regime nutrition

31. Tran, L.S., Nakashima, K., Sakuma, Y., *et al.* (2004). Isolation and functional analysis of Arabidopsis stressinducible NAC transcription factors that bind to a drought-responsive cis-element in the early responsive to dehydration stress 1 promoter, Plant Cell, 16, 2481–2498.
32. Vadivel V, Janardhanan K. Nutritional and antinutritional characteristics of seven South Indian wild legumes. Plant Foods for Human Nutrition. 2005 Jun;60:69-75.
33. Vandepoele, K., Quimbaya, M., Casneuf, T., De Veylder, L. and Van de Peer, Y.(2009). Unraveling transcriptional control in Arabidopsis using cis-regulatory elements and coexpression networks, Plant Physiol., 150, 535–46.
34. Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A. and Luscombe, N.M. (2009). A census of human transcription factors: function, expression and evolution, Nat. Rev. Genet., 10, 252–263.
35. Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA, Donoghue MT, Azam S, Fan G, Whaley AM, Farmer AD. Draft genome sequence of pigeonpea (Cajanuscajan), an orphan legume crop of resource-poor farmers. Nature biotechnology. 2012 Jan;30(1):83.
36. Vavilov, N.I. The origin, variation, immunity, and breeding of cultivated plants. *Chron. Bot.* 13, 1–366 (1951).
37. Wani, S., Tripathi, P., Zaid, A., Challa, G. S., Kumar, A., Kumar, V., Upadhyay, J., Joshi, R., Bhatt, M.(2018). Transcriptional regulation of osmotic stress tolerance in wheat (*Triticumaestivum* L.). Plant Molecular Biology, 97. 469-487.
38. Xiang, Y., Tang, N., Du, H., Ye, H. Xiong, L. (2008). Characterization of OsbZIP23 as a key player of the basic leucine zipper transcription factor family for conferring abscisic acid sensitivity and salinity and drought tolerance in rice. Plant Physiol. 148, 1938–1952.
39. Yamaguchi-Shinozaki, K. and Shinozaki, K. (2006). Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses, Annu. Rev. Plant Biol., 57, 781–803.
40. Zhang, W., Ruan, J., Ho, T.H., You, Y., Yu, T. and Quatrano, R.S.,(2005). Cis-regulatory element based targeted gene finding: genome-wide identification of abscisic acid- and abiotic stress-responsive genes in Arabidopsis thaliana, Bioinformatics, 21, 3074–81.