In-silico characterization of *ICE*1 gene of cold signalling pathway in mungbean (*Vigna radiata* L.)

Abstract

Aim: *In silico* characterization and identification of the *ICE1* cold tolerant gene pathway in mungbean *(Vigna radiata*), a proteinaceous legume is highly susceptible to cold stress.

Methodology: Genomic sequences of *Vigna radiata, V. angularis* and *V. unguculata and* other leguminous crops were retrieved from NCBI (https://pubmed.ncbi.nlm.nih.gov) and used for multiple sequence alignment, Phylogenetic and comparative analysis of ICE1 protein and motif analysis.

Results: The present study showed that *Vigna radiata ICE1 (VrICE1) gene* was closely related to *VaICE1* gene (*Vigna angularis)* and then *VuICE1* gene (*Vigna unguiculata)* . The study also revealed that the bHLH domain region (247-298aa) was the low complexity region and co-localization signals of *V.radiataICE1*. There were a similarities and dissimilarities of the ICE1 protein isoforms in *V. radiata*. It was noted that *VrICE*1 had an average of hydropathicity (GRAVY) of -0.593 in one and -0.6 in other. An instability index of 64.58 and 64.90. *VrICE*1 had a higher percentage of non-polar amino acid content and more number of random coils followed by alpha helics in secondary structure. The study also revealed that there was major involvement of *VrICE1* in biological and molecular functions under cold stress.

In conclusion, bioinformatics tools help researchers in getting information with regard to the functional aspect of the gene with respect to cold tolerance. A comprehensive analysis of the different physical and chemical properties of *Vigna* species could help us to identify their diversified usefulness.

Key words: *Vigna radiata*, *VrICE*1, bHLH, Physiochemical parameters, secondary structures.

1. Introduction

ICE1 is a MYC-type basic helix-loop-helix (bHLH) transcription factors that binding to the *cis*-elements in the promoter of C-repeat-binding factor3/dehydration-responsive element-binding protein1 A (*CBF3/DREB1A*) [1]. The *ICE*1 transcription factor, also known as *SCREAM (SCRM),* is anonymously considered as to be a master regulator of genes that

impart freezing tolerance [2,3,4,5]. *ICE1*, with simulation positively regulates the expression of *CBF3/DREB1A* in cold signalling. The *ICE1* mutation down-regulated the expression of *CBF3/DREB1A* and other *COR* gene decreased in both chilling tolerance and cold acclimation, whereas transgenic *ICE1*-overexpressing plants exhibited higher freezing tolerance [1]. Transcriptome analysis of the *ICE1* mutation is capable of down-regulating about 40 % of *COR* expression, accounting a huge share. Almost about 46 % of them were regulated by *ICE1* [6], when considering the cold regulated transcription factors. *ICE*2 was another MYC-type bHLH transcription factor with a high similarity to *ICE*1 and activates the expression of *CBF1/DREB1B* and is capable of promoting freezing tolerance in plants [7.8]. *ICE*1 is conserved in most of the plant species. Furthermore, the overexpression of *SlICE1* gene enhanced the accumulation of antioxidants, several amino acids, amines, and sugars and increased the antioxidant activity in tomato fruits [9]. *TaICE141* and *TaICE187* are also controlled the expression of the wheat *CBF* group IV genes and the overexpression of *TaICE187* and *TaICE141* in *Arabidopsis thaliana* augmented CBF/*DREB1*-dependent cold-responsive gene expression and cold acclimation [10]. Cold stresses induced the expression of the *OsICE*1 and *OsICE2* proteins, in rice and this in turn up regulated the expression of *OsDREB1B*, *OsHsfA3*, and *OsTPP1* suggesting that *OsICE1* and *OsICE2* function in the transcriptional regulation of cold response [11]. *MaICE1* interacted with *MaMYC2*, thereby triggered the activation of the jasmonate (JA) response pathway in fruit crop like banana [11, 12]. Furthermore, the expression of the cold response pathway genes was also induced by methyl Jasmonic acid (JA) significantly, suggesting that the JA-induced chilling tolerance of the banana plant is associated with induced expression of *MYC2*, which may act in coordination with *ICE1* for the activation of CBF-dependent cold signalling [12]. The overexpression of *Arabidopsis ICE1* also improved the chilling tolerance in *Vigna radiata* [13] and enhanced the accumulation of soluble sugars and proline in cucumber which moves in pace with the fact of the functional conservation of *ICE1* among species [14]. Many *in-silico* methods and online tools are available for the analysis and characterization of protein sequences. Different physico-chemical properties like protein length, amino acid composition, molecular weight, aliphatic index, extinction coefficient, isoelectric point, half-life, instability index and grand average of hydropathicity have also been reported [3,4,5]. There is no report on physicochemical properties and topology of various *ICE*1 genes of *Vigna radiate* so far. Hence, the present study was carried out to analyse and compare physicochemical properties and topology of ICE1 proteins of *Vigna radiata* in comparison to *Vigna unguiculata* and *Vigna angularis* and other leguminous

crops. This study will aid in understanding the occurrence of diversification in different ICE1 proteins of *Vigna radiata*. Further, it will also throw light on the similarities and differences among the protein sequences.

2. Materials and methods

2.1 Sequence retrieval

The sequence of ICE1 protein sequences for *Vigna radiata*, *V. angularis* and *V. unguculata* were retrieved from NCBI (https://pubmed.ncbi.nlm.nih.gov), in FASTA format and used for further analysis.

2.2 Multiple sequence alignment and phylogenetic and comparative analysis of ICE1 protein from *Vigna radiata* and other legume plants

Different legume crops sequences were retrieved from NCBI (https://pubmed.ncbi.nlm.nih.gov), using ICE1 from *Vigna radiata* as the query sequence. A pBLAST was performed and all the possible legume representative sequence was retrieved. The retrieved sequence was aligned using ClustalW and evolutionary relationship analysis of ICE1 protein using phylogenetic tree. Neighborjoining (NJ) method was used for the construction of phylogenetic tree of ICE1 in MEGA 11 software with bootstrap replication of 1000 times [15].

2.3 Physio-chemical properties

The physicochemical properties were computed for all the proteins using the ExPASy ProtParam tool (http://web.expasy.org/protparam/) [16].Web servers specialized in predicting cellular localization of protein sequence were analysed using WoLF PSORT (http://wolfpsort.seq.cbrc.jp/) [17] for theoretical measurements, such as molecular weight, isoelectric point (pI), extinction coefficient, instability index, and aliphatic index. Functional domain, their subcellular localization and presence of transmembrane domain of different ICE1 proteins of *Vigna* was assessed. The structure of the domain was created using SMART tool (http://smart.embl-heidelberg.de/)

2.4. DNA binding logo and motif analysis

Sequences of *Vigna radiata* ICE1 (VrICE1) proteins were aligned using the WEBLOGO 3 program version 2.8.2 at http://weblogo.berkeley.edu/ to get the DNA binding logo [18]. The conserved motifs of ICE1 proteins were assessed via MEME tools (http://meme.sdsc.edu/meme/intro.html).

2.4. Secondary structure prediction

Secondary structures of ICE1 proteins were predicted with online Expasy SOPMA tool (Self-Optimized Prediction Method and Alignment) (https://npsaprabi.ibcp.fr/cgibin/npsa_automat.pl?page=/NPSA/npsa_sopma.html) **[19].** This tool caters to information regarding about different conformations of proteins such as percentages of α-helices, β-sheets, turns, extended strands and random coils. The Self-Optimized Prediction method With Alignment (SOPMA) tool is used to predict the secondary structure of a protein. Based on the query (primary sequence of a protein), SOPMA will predict its secondary structure.

2.6 .Protein- protein interaction network

The ICE1 protein interaction network and co-expression analysis was assessed using STRING (https://string-db.org/) database. The interactions analyzed include direct (physical) and indirect (functional) association and computational prediction, knowledge transfer between organisms, and interactions aggregated from of other (primary) databases. The gene ontology was also studied using STRING database gene ontology function (Add ref.).

3. Results and Discussion

3.1. Phylogenetic analysis of ICE1 protein of *Vigna radiata* and other legume plants.

The phylogenetic analysis reveals that they are mainly divided into 2 main clades. Of which VrICE1 belongs to the smallest 24 member sub-clades. Among the sub-clades of *Vigna sp., Vigna angularis* is significantly more similar to *V radiata* than *V. unguiculata.* This was evident from the phylogentic tree (**Fig.1**). The next set of legumes i.e. *Glycine max* that are very closely related to VrICE1 protein and followed by *Cajanus cajan*. The largest group is 13 member clade, but that is distantly related to VrICE1 proteins.

3.2. Domain analysis and Co-localization of protein

The gene structure reveals the presence of bHLH domain in the ICE1 protein of *Vigna*. The bHLH domain was found in the region from 247-296 amino acids (aa) as indicated in the **Table-1**. This was similar with all *Vigna* species, whereas it ranged from 249-298aa in *V.unguiculata (Vu)* and 242-291aa and 204-253aa in *Vigna angularis (Va).* VaICE1 of was predicted to have a transmembrane domain, which was not found in VrICE1. The spread of low complexity region is same in case of both VrICE1 and VaICE1 proteins. The protein sequences showed sub-cellular localization signals for nuclear and cytoplasmic localization in case of *V.radiata*. It also showed that ICE1 localization signal observed in nucleus, cytoplasm, vacuoles, golgibodies and chloroplast of *V.unguiculata* and *V.angularis* (**Table 1**).

3.3 Analysis of Physio-chemical properties

Protein length, isoelectric point (pI), molecular weight, extinction coefficient, total number of negatively and positively charged residues, instability index (II), aliphatic index (AI) and grand average of hydropathicity (GRAVY) are the physiochemical properties taken into account for VrICE1. *Vigna sp* were also computed using ExPASy ProtParam tool. Subcellular localization, solubility and interactions dependent on the parameters like number of positively and negatively charged residues and isoelectric point. The physiochemical parameters show that numbers of the amino acid ranged from 391aa to 455 aa in VrICE1 and the molecular weight of 47898.22 Da and 48026.35 Da. The isoelectric point ranged from 5.26 to 6.00 which among similar with the *Vigna* sp. This shows the acidic nature of VrICE1. Protein EC value was an important parameter based on amount of light absorbed per mole of protein at a 280 nm wavelength. Extinction coefficient (EC) value of protein is derived from the number of tyrosine, tryptophan and cysteine residues per molecule. This is because these residues contribute significantly to measured optical density of denatured protein at 276–282 nm range [20,21]. All the studied *Vigna sp*. shared the same EC 2168M^{-1} cm^{-1} at 280 nm, but VaICE1 had 20315 and 14690 M^{-1} cm^{-1} at 280 nm. The low value was due to low amount of tryptophan and tyrosine residues in the amino acids. Instability index (II) is an indication of the stability of a protein under both *in-vitro and in-vivo* conditions. Proteins with instability index (II) <40 are considered to be stable and those with II value >40 are named to as unstable [22,23]. VrICE1 showed the highest 64.90 instability index (II) which fall in unstable category. Aliphatic index (AI) is another parameter to check the stability of the proteins. AI of a protein can be explained as the relative volume captured by aliphatic side chains of amino acids like A (alanine), V (valine), L (leucine) and I (isoleucine). Thermo-stability and AI of proteins were highly correlated with *Vigna radiata* [4,23]. Further, VrICE1 had an aliphatic Index of 71.04 and 70.87, falling in an average range of thermostability among the species under study. The range of all the *Vigna sp*. studied had fallen to be in 68.28 to 72.10.

Apart from studying protein concentration and stability, its hydrophilic character or hydrophobic character was also analyzed with the GRAVY score [19]. GRAVY score for particular protein was the sum of hydropathy values of all amino acids present in the protein, divided by the number of residues in that protein. The values ranges between -2 to +2 where; negative score means hydrophilicity and positive score indicates hydrophobicity [19, 24, 27]. Proteins with more to be hydrophilic in nature with good solubility and vice-versa, if they have negative GRAVY. If a protein has GRAVY score more than 0.4, suggest its

hydrophobic nature and difficult to detect on 2-D gel electrophoresis [24] with a GRAVY range of -0.571 to -0.609, Vr*ICE*1 has -0.593 and -0.600 (**Table 2.),** which states that it is hydrophilic in nature. Considering the amino acids present in *VrICE*1, it has been observed that there is a higher percentage of non-polar amino acid, particularly leucine and proline. Proline has been shown to act like a molecular chaperon who provides protection against abiotic and biotic stresses by enhancing activities of some enzymes as well as maintaining integrity of proteins [25]. All the studied species had very low amount of aromatic amino acids. *VuICE*1 showed very less methionine compared to the other sp. and histidine percent was very less in all the *Vigna sp*. (**Fig 2**).

3.4 Motif analysis and DNA binding Logo

There were 10 prominent motifs detected and is named 1-10 (**Fig 3**). The motif 1, was found to be the most prominent motif among them, both in case of size and conservation. Motif 1 to 7 was found in all the sequences studied. Motif 8 was absent in VuICE1. Between the VrICE1, Motif 10, which is the shortest motif (**Fig 4**) is the sole difference, which is absent in the VrICE1 isoform. It was found to be common in all the *Vigna* ICE1 protein (**Fig 3).** The conserved DNA binding region of the VrICE1 protein had been used for logo. It shows that the conserved region ranges from 1-35aa and from 110 to 131 aa (in order of amino acid arrangement in ICE1 protein), among the *Vigna sp*. studied (**Fig 5)**

3.5 Secondary structure prediction

The secondary structure predicted the abundance of secondary structures in the *Vigna sp ICE1* genes. The number followed the same pattern of random coils > alpha helices>extended strands > beta turns in all the *Vigna sp*. under study without much significant difference, indicating that there is no much difference in the secondary structural patterns in the genes (**Fig 6**). But on a detailed observation on the arrangement patterns of these secondary structures, it was seen that they differ in their arrangement significantly from each other due to the difference in the arrangement pattern, number and abundance of amino acids present in *Vigna* ICE1 protein (**Fig 7**). In case of *VrICE*1 genes, the random coils were covering 60.14 and 57.75%, the alpha helix was in 27.4 and 29.43 %. This was followed by extended strand which was 9.68 % and 8.97%. The least was covered by beta turns which were 2.76% in both VrICE1 and VuICE1. VuICE1 showed a relatively higher amount of alpha helix (31.49%), but had the least amount of random coils (56.55%) among the *Vigna sp*. studied.

3.6. Protein- protein interaction network analysis of VrICE1

The protein interaction network analysis showed that *VrICE*1 has direct interaction with several genes, of which the prominent ones are FAMA, HOS1, heptahelical transmembrane proteins, MAPKase, MUTE and E3 SUMO SIZ1 protein ligase. The prominence of interaction was indicated by the prediction score of 0.903, 0.896, 0.885, 0.883, 0.871 and 0.838, respectively. FAMA is basic helix-loop-helix (bHLH) DNA-binding super family protein, a transcription activator. Together with MYB88 and MYB124, ensures that stomata contain two guard cells by enforcing a single symmetric precursor 56 cell division before stomatal maturity. Together with *SPCH* and *MUTE*, regulates the stomata formation. It is required to promote differentiation and morphogenesis of stomatal guard cells and to halt proliferative divisions in their immediate precursors. It mediates the formation of stomata [26] and prevents histone H3K27me3 marks and depress the stem cell gene expression FAMA gene having a co-expressional score of 0.1 (Fig 8). It was found that cold stress triggers protein degradation of nuclear GFPICE1 protein, and the RING finger protein. Sumoylation of *ICE*1 controls CBF3/DREB1A expression and freezing tolerance. This was also evident that *CE1/SCRM* was suggested to be involved in stomatal differentiation **[27]** and *ICE1/SCRM* and *ICE2/SCRM2* are paralogous genes. The *ice1-2 scrm-2* double mutant exhibited lack of stomata differentiation from epidermal cells, whereas *ice1-1*, a dominant negative mutant, displayed many stomata [25]. So, it is that *ICE1* may be involved in integration of cold signalling and stomatal development when it interacts with SPCH, MUTE, and FAMA [21].

In fig 9, it shows the co-expression scores of the indicated genes namely, are FAMA, HOS1, heptahelical transmembrane proteins, MAPKase, MUTE and E3 SUMO SIZ1 protein ligase in other species of the plant kingdom. High similarity in patterns of expression of the linked gene was observed in *Glycine sp*. (Soybean) and *Phaseolus sp*, in comparison to *Vigna*. This indicates its less divergence from the common ancestry patterns in gene expression. Few gene expressions are observed in fungi too, tracing the evolutionary conservation in eukaryotes, which was absent in prokaryotes.

3.7 Gene ontology of Vr ICE1 protein

Gene ontology of Vr ICE1 was capable of showing its importance in the molecular function of MAPKase and SUMO transferase activity (**Table 3b)** and the biological functions of response to freezing and several other associated functions. Among the other functions, response to nitrate assimilation, detection of phosphate ion, regulation of systemic acquired

resistance, response to freezing, pollen tube guidance and stomatal lineage progressions are the top scored functions of VrICE1. This was followed by salicylic acid function, ovule development, heat acclimation; regulation of response to water deprivation, regulation of ABA regulated pathways and gametophyte development. Thus, VrICE1 protein a major focus of the study in cold tolerance. (**Table 3a**).

Conclusion

In conclusion, *in-silico* characterization of ICE1 protein in *V.radiata* was performed and compared with other *Vigna* species and other leguminous crops. The analysis of ICE1 was performed, including phylogeny, motif identification, physiochemical characterization, colocalization analysis, protein regulatory network, co-expression analysis and gene ontology analysis. The results indicated that the involvement of *ICE*1 transcription factors in a multitude of biological and molecular function. The study throws light on the variants in the central properties such as molecular weight, isoelectric point, and total number of negatively, positively charged residues, extinction coefficient, instability index, aliphatic index and grand average of hydropathicity within *ICE1* proteins, which may be responsible for their functional diversity. Although significant progress has been made to understand and identify the molecular mechanism of cold acclimation and the cold sensor (s) has not yet been identified. Thus, further studies for the identification of the cold sensor(s) will be necessary to understand how plants sense temperature.

Data availability statement

Data will available with the department. On request it will be provided.

References

1. Chinnusamy V, OhtaM., Kanrar S., Lee BH, Hong X, Agarwal, M, and ZhuJ.K. *ICE*1: A regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis.* Genes Dev. 2003;17: 1043–1054.

2. Yamaguchi-Shinozaki K., and Shinozaki K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. Annu. Rev. Plant Biol. 2006;57: 781–803.

3. Knight, M.R., and Knight, H. Low temperature perception leading to gene expression and cold tolerance in higher plants. New Phytol. 2012;195: 737–751.

4.Guo, X., Liu, D., and Chong, K. Cold signalling in plants: Insights into mechanisms and regulation. J. Integr. Plant Biol. 2018;60: 745–756.

5.Shi Y., Ding, Y., and Yang, S. Molecular regulation of CBF signaling in cold acclimation. Trends Plant Sci. 2018; 23: 623–637.

6. Lee BH, Henderson DA, Zhu JK The Arabidopsis cold-responsive transcriptome and its regulation by ICE1. Plant Cell 2005; 17:3155–3175.

7.Fursova OV, Pogorelko GV, Tarasov VA Identification of ICE2, a gene involved in cold acclimation which determines freezing tolerance in Arabidopsis thaliana. Gene 2009; 429:98–103.

8.Miura K, Shiba H, Ohta M, Kang SW, Sato A, Yuasa T, Iwaya-Inoue M, Kamada H, Ezura H (2012a) SlICE1 encoding a MYC-type transcription factor controls cold tolerance in tomato, Solanum lycopersicum. Plant Biotechnol 29:253–260.

9.Miura K, Sato A, Shiba H, Kang SW, Kamada H, Ezura H (2012b) Accumulation of antioxidants and antioxidant activity in tomato, Solanum lycopersicum, are enhanced by the transcription factor SlICE1. Plant Biotechnol 29:261–269.

10 Badawi M, Reddy YV, Agharbaoui Z, Tominaga Y, Danyluk J, Sarhan F, Houde M (2008) Structure and functional analysis of wheat ICE (inducer of CBF expression) genes. Plant Cell Physiol 49:1237–1249

11. Nakamura J, Yuasa T, Huong TT, Harano K, Tanaka S, Iwata T, Phan T, Iwaya-Inoue M (2011) Rice homologs of inducer of CBF expression (OsICE) are involved in cold acclimation. Plant Biotechnol 28:303–309

12. Zhao ML, Wang JN, Shan W, Fan JG, Kuang JF, Wu KQ, Li XP, Chen WX, He FY, Chen JY, Lu WJ (2012) Induction of jasmonate signalling regulators MaMYC2s and their physical interactions with *MaICE*1 in methyl jasmonate-induced chilling tolerance in banana fruit. Plant Cell Environ 36(1):30–51

13. Rout G, Bansal A , Swain, D Jadhao K , Shelke, R and Panda, S. Overexpression of *ICE1* gene in mungbean (*Vigna radiata* L.) for cold tolerance 2020;143: 593-608

14. Liu L, Duan L, Zhang J, Zhang Z, Mi G, Ren H (*Cucumber (Cucumis sativus L.*) over-expressing cold-induced Transcriptome regulator ICE1 exhibits changed morphological characters and enhances chilling tolerance. Sci Hortic 2010; 124:29–33

15. Kumar, S., Stecher, G., and Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 2006;33, 1870–1874.

16. Gasteiger E,.The proteomics protocols handbook. 2005;571.

17. Srinivas V R. Bioinformatics - A Practical Guide to the analysis of Genes and Proteins. 2005, ISBN : 978-81-203-2858-7, published by PHI Learning Private Limited, New Delhi

18 .Timothy L. Bailey, Charles Elkan, Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, 1994 pp. 28-36, AAAI Press, Menlo Park, California.

19. Geourjon C, Deléage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics.1995;11:681–684.

20. Edelhoch H. Spectroscopic Determination of Tryptophan and Tyrosine in Proteins. Biochemistry. 1967; 6: 1948–1954.

21. B.Kaur G, Pati PK. *In silico* physicochemical characterization and topology analysis of respiratory burst oxidase homolog (Rboh) proteins from *Arabidopsis* and *rice*. Bioinformation. 2018;14:93–100.

22.Kenji Miura, 2013, ICE1, A transcription factor involved in cold signalling and tolerance plant and microbe adaptations to cold in a changing world, 2012; 978-1-4614-8252-9.

23.Guruprasad K, Reddy BVB, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. Protein Engineering. 1990; 4: 155–161.

24. Wilkins MR, Gasteiger E, Wheeler CH, Lindskog I, Sanchez JC, Bairoch A, Two-dimensional gel electrophoresis for proteome projects: the effects of protein hydrophobicity and copy number. Electrophoresis. 1998; 19:1501–1505.

25. Szabados L and Savour A. Proline: a multifunctional amino acid. Trends in Plant Science. 2009; 15: 89–97.

26. Ikai A. Thermostability and aliphatic index of globular proteins. Journal of Biochemistry. 1980;88:1895–98.

27. Kanaoka MM, Pillittere LJ, Fujii H, Yoshida Y, Bogenschutz NL, Takabayashi J, Zhu JK, Torii KU SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to Arabidopsis stomatal differentiation. Plant Cell . 2008; 20:1775–1785.

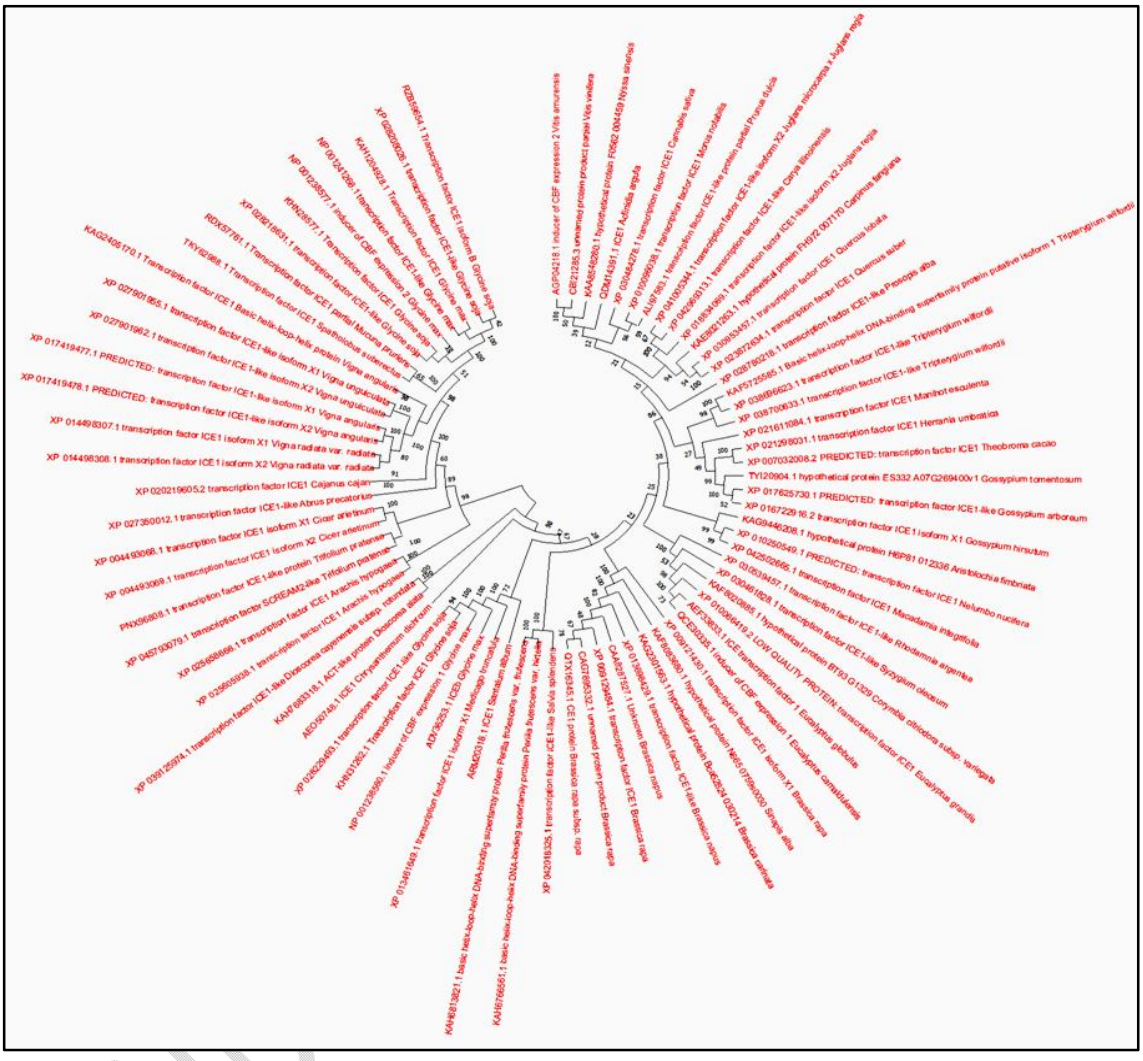
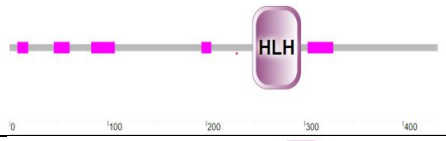
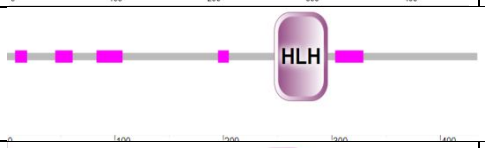
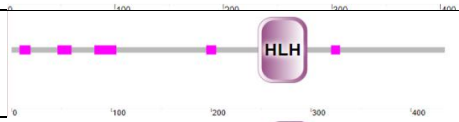
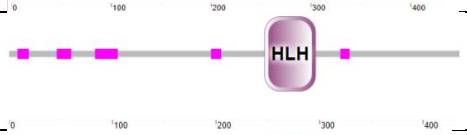
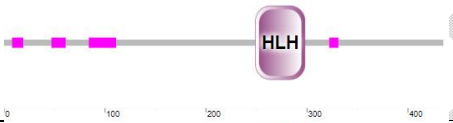
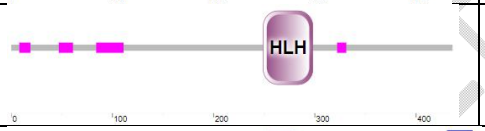
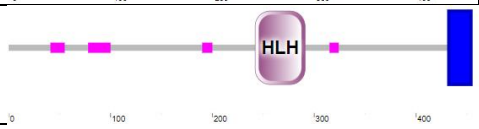
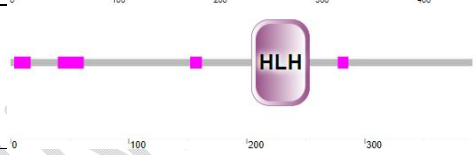


Fig 1. ICE1 proteins of legumes were used to create the phylogenetic tree using MEGA11 with Neighbourhood joining method. Evolutionary relationship analysis of ICE1 proteins with other *Vigna s*pecies.

Table 1. Functional domain, their subcellular localization and presence of transmembrane domain of different ICE1 proteins of *Vigna spp.*

Accession number	DOMAIN STRUCTURE	HLH Region		Low complexity region		Subcellular localization Signals	Transmembrane domain region
		Start - End		Start	End		Start End
>XP_014498307.1 transcription factor ICE1 isoform X1 [*Vigna radiata var. radiata*]		247	296	8 45 83 195 303	19 61 107 205 329	**Nuclear**/Cytoplasmic	N/A
>XP_014498308.1 transcription factor ICE1 isoform X2 [*Vigna radiata var. radiata*]		247	296	8 46 83 195 320	19 60 105 205 329	**Nuclear**/Cytoplasmic	N/A
>XP_017419478.1 PREDICTED: transcription factor ICE1-like isoform X2 [*Vigna angularis*]		247	296	8 46 83 195 320	19 60 105 205 329	**Nuclear**/Cytoplasmic	N/A
>XP_017419477.1 PREDICTED: transcription factor ICE1-like isoform X1 [*Vigna angularis*]		247	296	8 46 83 195 320	19 60 105 205 329	**Nuclear**/Cytoplasmic	N/A
>XP_027901962.1 transcription factor ICE1-like isoform X2 [*Vigna unguiculata*]		249	298	8 47 84 322	19 61 111 331	**Nuclear**/Cytoplasmic /Cytoskeleton	N/A
>XP_027901955.1 transcription factor ICE1-like isoform X1 [*Vigna unguiculata*]		249	298	8 47 84 322	19 61 111 331	**Nuclear**/Cytoplasmic /Cytoskeleton	N/A
>KAG2405170.1 Transcription factor ICE1 Basic helix-loop-helix protein [*Vigna angularis*]		242	291	41 78 190 315	55 100 200 324	**Nuclear**/Cytoplasmic/ Vacuolar/golgi/ chloroplast	432 to 454
>KOM38329.1 hypothetical protein LR48_Vigan03g171100 [*Vigna angularis*]		204	253	3 40 152 277	17 62 162 286	**Nuclear**/ Cytoplasmic and nuclear/ ER	N/A

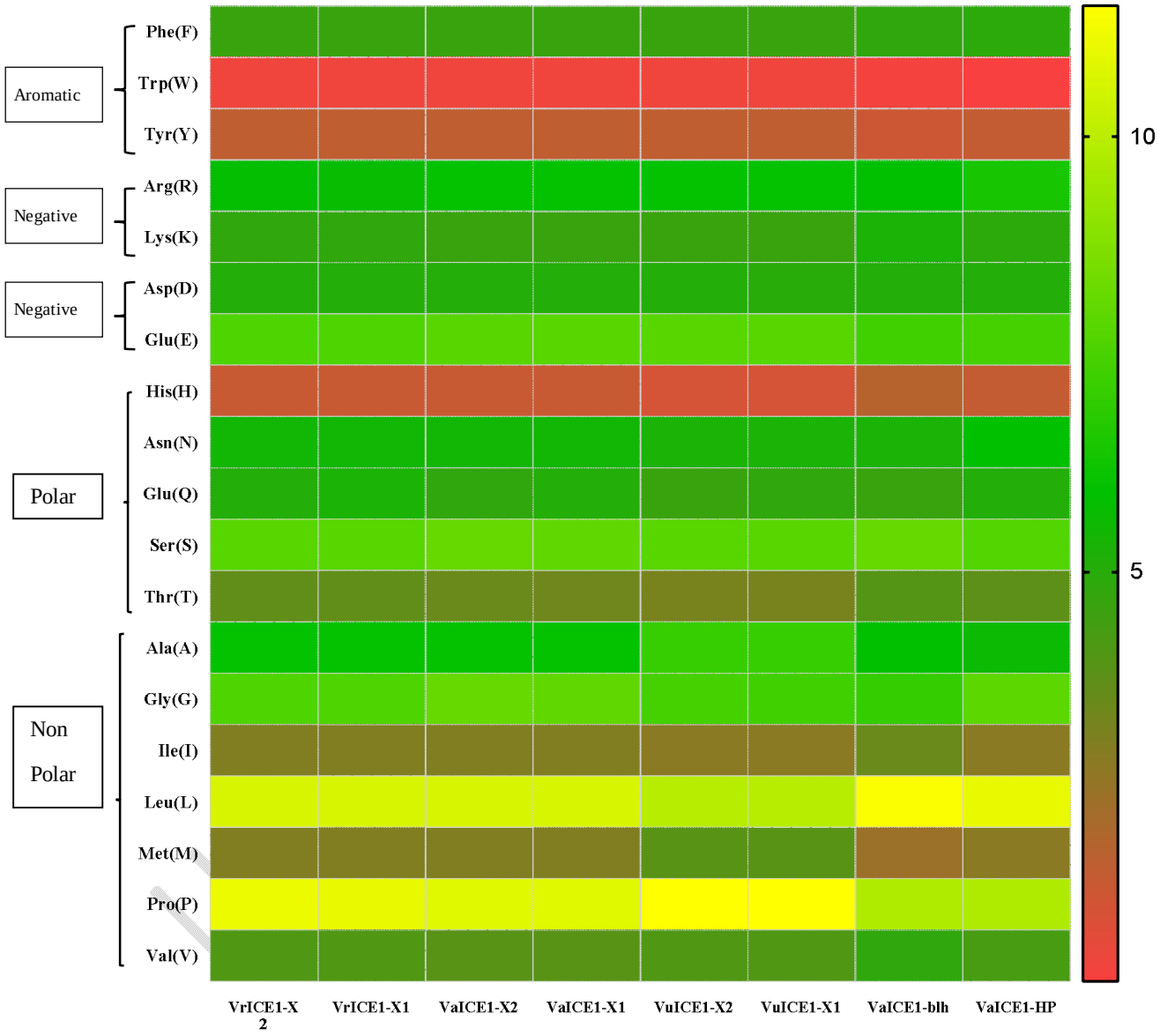
.

Table 2. Physico-chemical properties of ICE1 proteins of *Vigna* computed using ProtParam tool.

Accession number	Amino acid length	M.Wt	pI	(-)R	(+)R	Ec[a]	II	AI	GRAVY	Mol. Formulae and Atomic number
>XP_014498307.1 transcription factor ICE1 isoform X1 [Vigna radiata var. radiata]	434	47898.22	5.40	55	46	21680 21430	64.58, unstable	71.04	-0.593	$C_{2110}H_{3320}N_{590}O_{648}S_{18}$ (6686)
>XP_014498308.1 transcription factor ICE1 isoform X2 [Vigna radiata var. radiata]	435	48026.35	5.40	55	46	21680 21430	64.90 unstable	70.87	-0.600	$C_{2115}H_{3328}N_{592}O_{650}S_{18}$ (6703)
>XP_017419478.1 PREDICTED: transcription factor ICE1-like isoform X2 [Vigna angularis]	434	47831.04	5.30	56	46	21680 21430	63.54 unstable	70.37	-0.603	$C_{2110}H_{3320}N_{590}O_{648}S_{18}$ (6668)
>XP_017419477.1 PREDICTED: transcription factor ICE1-like isoform X1 [Vigna angularis]	435	47959.17	5.33	56	46	21680 21430	63.86 unstable	70.21	-0.609	$C_{2108}H_{3315}N_{593}O_{651}S_{18}$ (6685)
>XP_027901962.1 transcription factor ICE1-like isoform X2 [Vigna unguiculata]	435	47910.34	5.26	56	46	21680 21430	63.68 unstable	68.44	-0.571	$C_{2106}H_{3312}N_{588}O_{646}S_{22}$ (6674)
>XP_027901955.1 transcription factor ICE1-like isoform X1 [Vigna unguiculata]	436	48038.47	5.26	56	46	21680 21430	64.00 unstable	68.28	-0.578	$C_{2111}H_{3320}N_{590}O_{648}S_{22}$ (6691)
>KAG2405170.1 Transcription factor ICE1 Basic helix-loop-helix protein [Vigna angularis]	455	50332.25	6.00	56	51	20315 19940	58.93 unstable	78.24	-0.486	$C_{2225}H_{3525}N_{623}O_{674}S_{17}$ (7064)
>KOM38329.1 hypothetical protein LR48_Vigan03g171100 [Vigna angularis]	391	43281.98	5.82	49	44	14690 14440	62.20 unstable	72.10	-0.598	$C_{1900}H_{3003}N_{541}O_{584}S_{16}$ (6044)

M.wt. = Molecular weight; pI = Isoelectric point; (-) R = total number of negatively charged residues; (+) R = total number of positively charged residues; Ec[a] = Extinction coefficient (a units of M^{-1} cm^{-1}at 280 nm measuring in water); II = Instability index; AI = Aliphatic index; GRAVY = Grand average of hydropathicity. The N-terminal of the sequence considered is M (Met). The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro), >20 hours (yeast, in vivo), >10 hours (Escherichia coli, in vivo) for all the proteins

Fig 2. Heatmap showing the comparison of Amino acid composition of *ICE*1 Proteins of *Vigna spp*.



VrICE1-X1 indicates XP_014498308.1 transcription factor ICE1 isoform X2 [*Vigna radiata var. radiata*]

VrICE-X2 indicates XP_014498307.1 transcription factor ICE1 isoform X1 [*Vigna radiata var. radiata*]

VaICE1-X2 indicates XP_017419478.1 PREDICTED: transcription factor ICE1-like isoform X2 [*Vigna angularis*]

VaICE1-X1 indicates XP_017419477.1 PREDICTED: transcription factor ICE1-like isoform X1 [*Vigna angularis*]

VuICE-X2 indicates XP_027901962.1 transcription factor ICE1-like isoform X2 [*Vigna unguiculata*]

VuICE1-X1 indicates XP_027901955.1 transcription factor ICE1-like isoform X1 [*Vigna unguiculata*]

VaICE1-BLH indicates KAG2405170.1 Transcription factor ICE1 Basic helix-loop-helix protein [*Vigna angularis*]

VaICE1-HP indicates KOM38329.1 hypothetical protein LR48_Vigan03g171100 [*Vigna angularis*]

Fig 3. Motifs of the VrICE1 proteins is identified by MEME.The motifs numbered 1 to 10, are displayed in different colours.

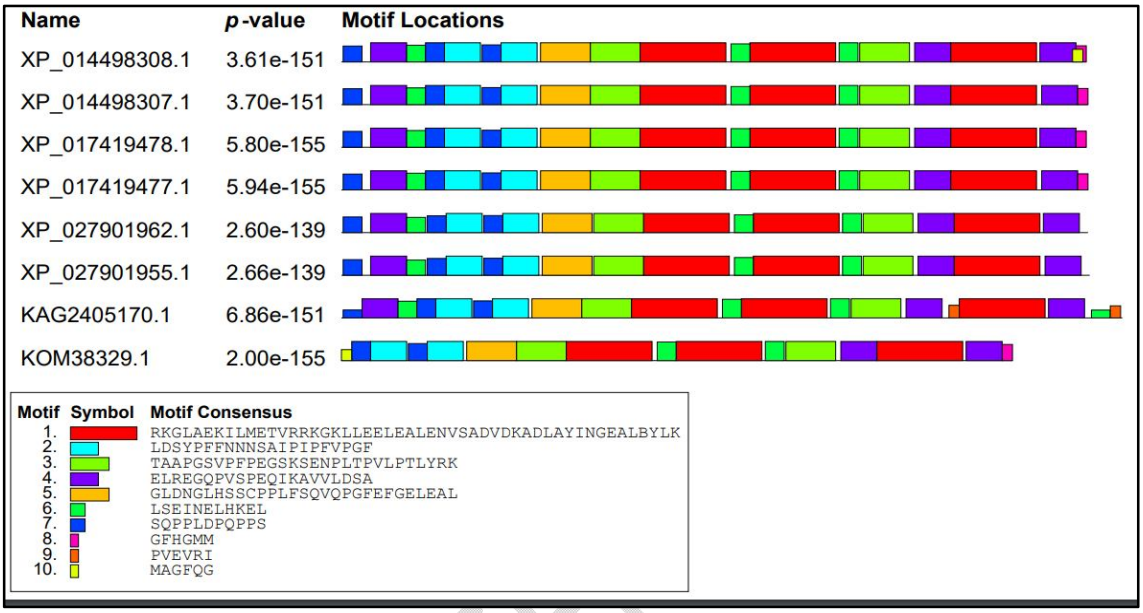


Fig4. Motifs logos of the *ICE*1 proteins. The “sites” indicate the number of Vr*ICE*1 proteins containing each motif. The “width” indicates the amino acid number of each motif.

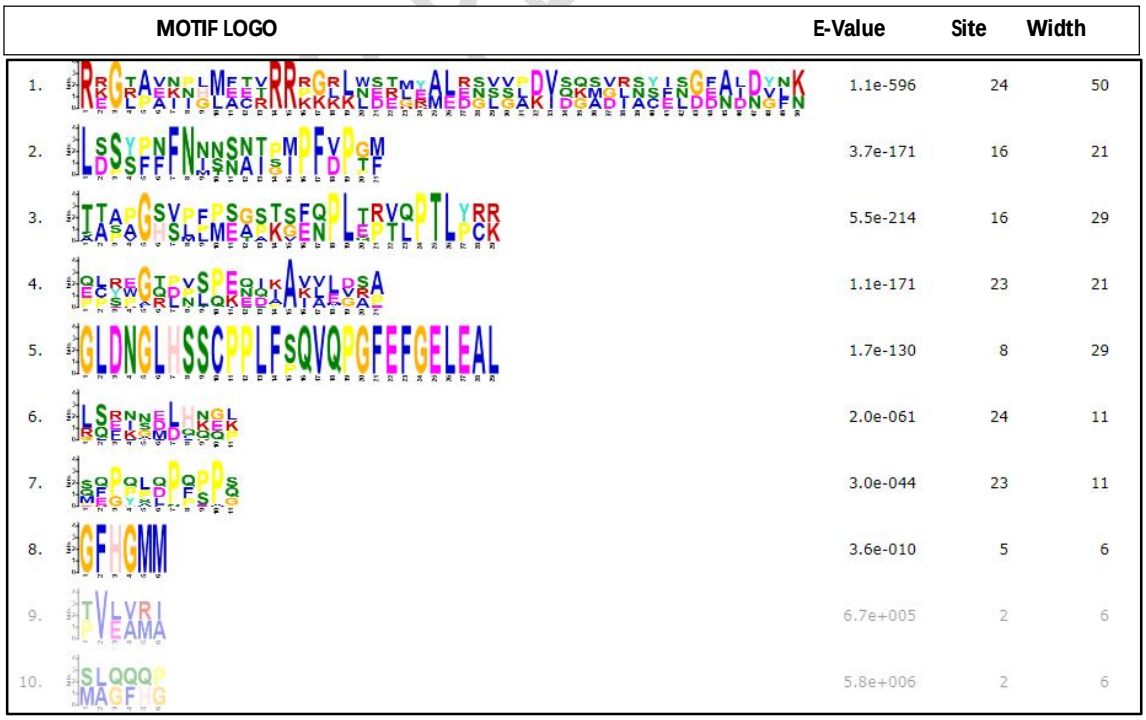


Fig 5. DNA Binding domain alignment of *ICE*1 protein from *Vigna radiata*.

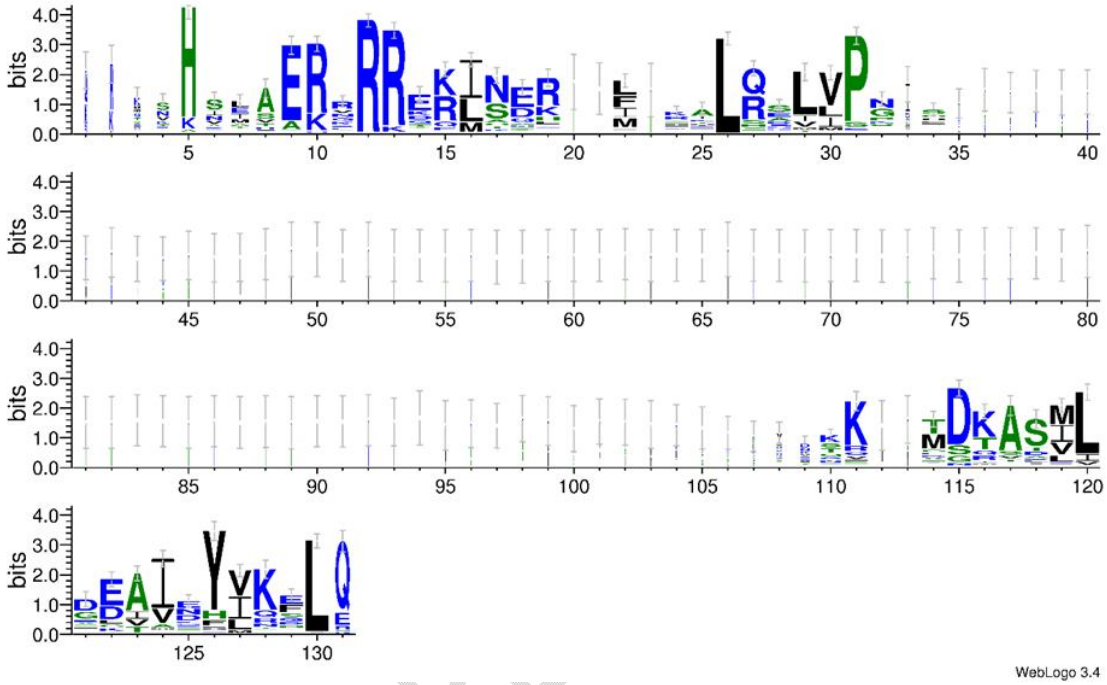
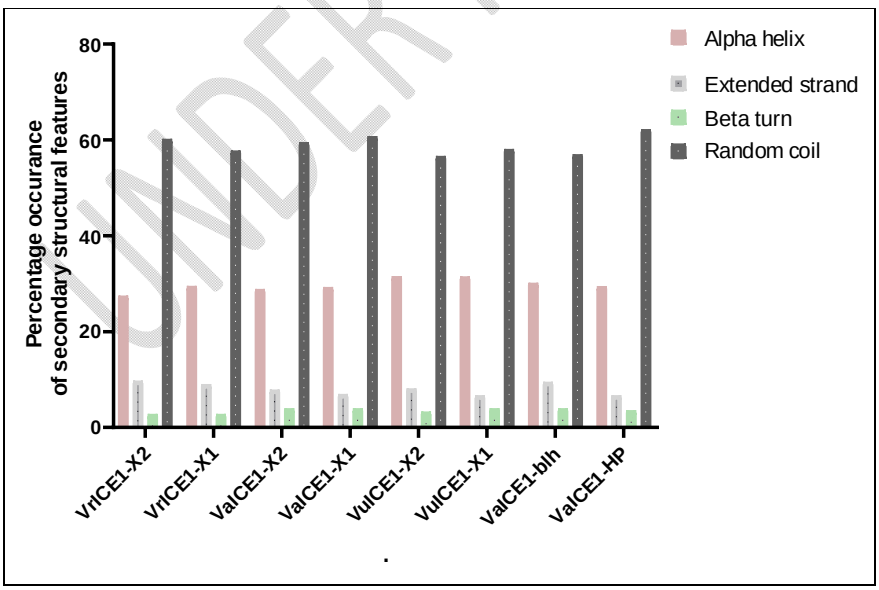


Fig 6. Percentage occurrence of secondary structural features of ICE1 protein in *Vigna* spp.

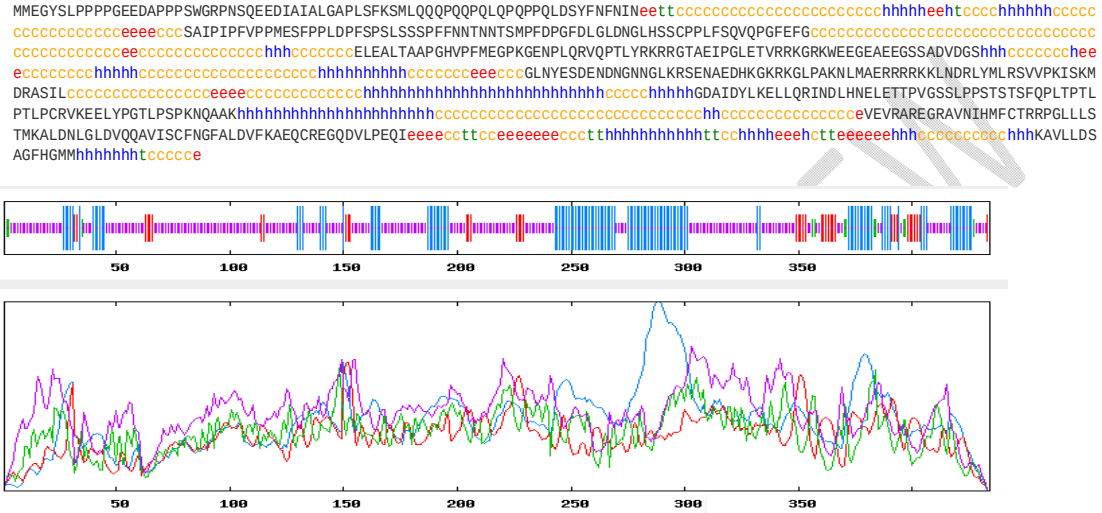


VrICE1-X1 indicates XP_014498308.1 transcription factor ICE1 isoform X2 [*Vigna radiata var. radiata*]
VrICE-X2 indicates XP_014498307.1 transcription factor ICE1 isoform X1 [*Vigna radiata var. radiata*]
VaICE1-X2 indicates XP_017419478.1 PREDICTED: transcription factor ICE1-like isoform X2 [*Vigna angularis*]

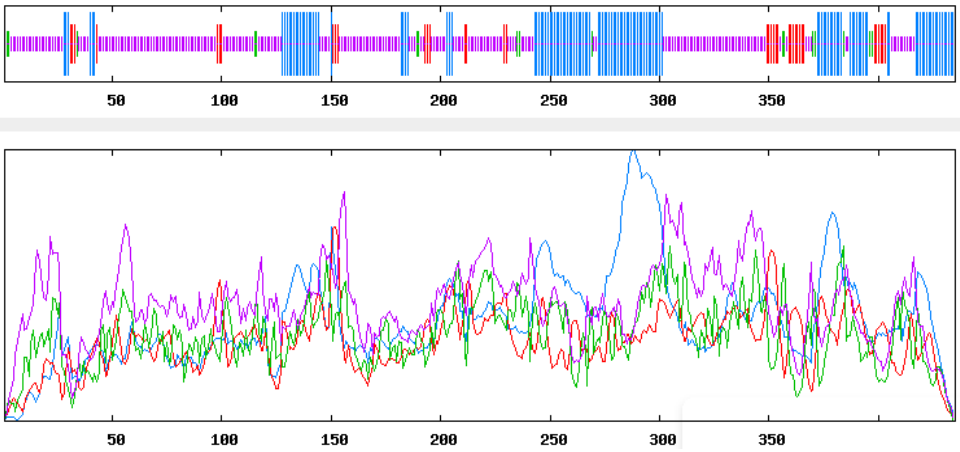
VaICE1-X1 indicates XP_017419477.1 PREDICTED: transcription factor ICE1-like isoform X1 [*Vigna angularis*]
VuICE-X2 indicates XP_027901962.1 transcription factor ICE1-like isoform X2 [*Vigna unguiculata*]
VuICE1-X1 indicates XP_027901955.1 transcription factor ICE1-like isoform X1 [*Vigna unguiculata*]
VaICE1-BLH indicates KAG2405170.1 Transcription factor ICE1 Basic helix-loop-helix protein [*Vigna angularis*]
VaICE1-HP indicates KOM38329.1 hypothetical protein LR48_Vigan03g171100 [*Vigna angularis*]

Fig 7. Detailed representation of all the predicted secondary structures in ICE1 proteins in *Vigna radiata, V. angularis* and *V. unguculata.*

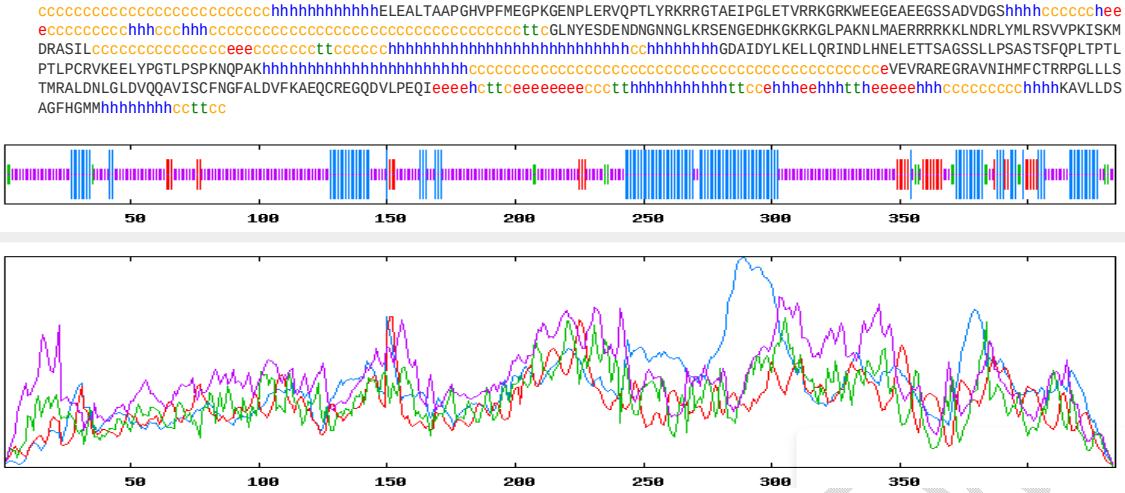
1. XP_014498308.1 transcription factor ICE1 isoform X2 [Vigna radiata var. radiata]



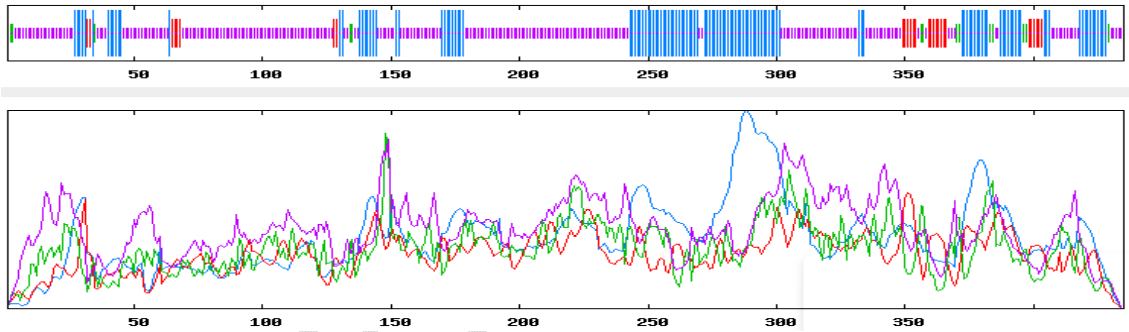
2. **XP_014498307.1 transcription factor ICE1 isoform X1 [Vigna radiata var. radiata]**



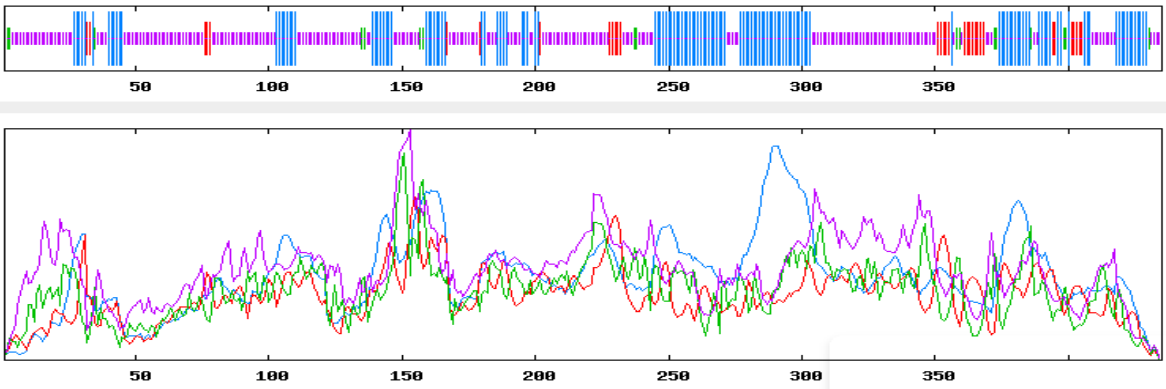
3. XP_017419478.1 PREDICTED: transcription factor ICE1-like isoform X2 [Vigna angularis]



4. XP_017419477.1 PREDICTED: transcription factor ICE1-like isoform X1 [Vigna angularis]

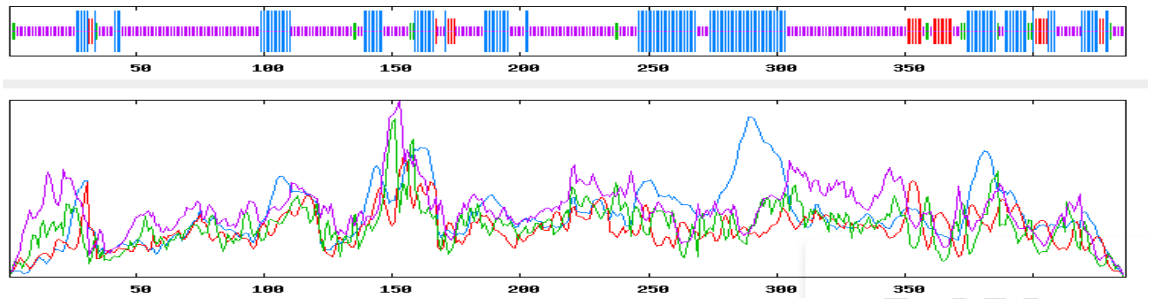


5. XP_027901962.1 transcription factor ICE1-like isoform X2 [Vigna unguiculata]



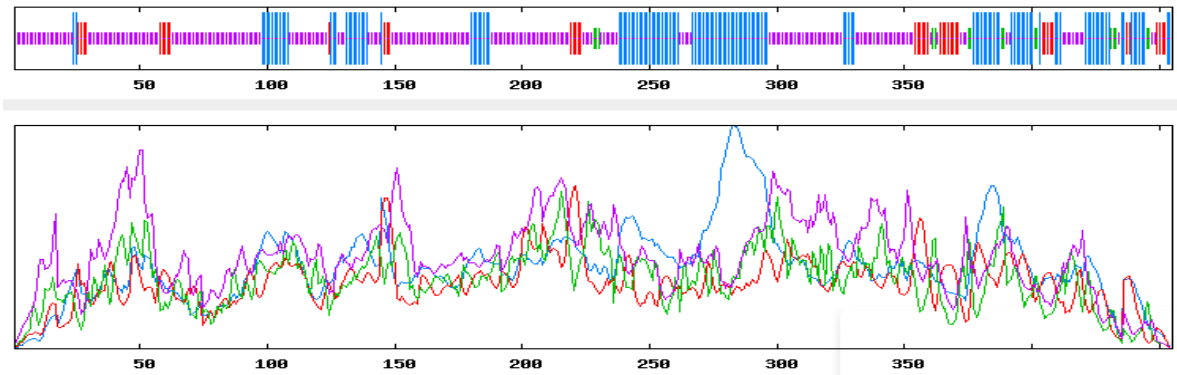
6. VuICE1-X1 indicates XP_027901955.1 transcription factor ICE1-like isoform X1 [Vigna unguiculata]

MMEGYALPPPPGEEDAPPPSWGRPNSQEEDIAIALGAPLSFKSMLLQQPQPQPQPQPQPPQLDPYFNFNMeettccccccccccccccccccccccchhhhheehtccccccchhhccccccccccc
ccccccccccccccNSAIPIPFVPTMESFPPLDPFSPSLSSSPFFNNSSNTSMPFDPGFDMGLDNGLHSSCPPLFPQVQPGFEFcccccccccccccccccccccccccccchhhhhhhhhhhhc
ccccccccccccccccccccccttcchGELEALTAAAPGHVPFMEGAKVENPLERVQPTLYRKRRGTAEIPGLETVRRKGRKWEEREAEEGSSADVDhhhhhhhccccccccctthhhhhhhhe
cchheeecccccccccchhhhhhhhhhccccchhccccccGSGLNYESDENDNGNNGLKGSENGEDHKGKRKGLPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRAScccccccccccccc
ccccccccccccttccccccchhhhhhhhhhhhhhhhhhhhhhhcccchhhhhhILGDAIDYLKELLQRINDLHNELESTPAGSSMPPSASTSFQPLTPTLPTLPCRVKEELYPATLPSPKNQA
hhhhhhhhhhhhhhhhhhhhhhhhcccccccccccccccccccccccccccccccccccccccccccccccccAKVEVRAREGRAVNIHMFCTRRPGLLLSTMRALDNLGLDVQQAVISCFNGFALDVF
KAEQQCREGQDVLPceeeeeecttceeeeeeeeccctthhhhhhhhhhhhtccchhhhhhhhhttheeeeehhhccccccccchEQIKAVLMDSAGFGMMhhhhhheehhtccccc



7. VaICE1-BLH indicates KAG2405170.1 Transcription factor ICE1 Basic helix-loop-helix protein [Vigna angularis]

MDKPHPFNQDAPPPSWGRPNSQEEDIAIALGAPLSFKSMLQQQPQQPQPQPQPPQLDSYFNFNINSAIPIccccccccccccccccccccccchheeeecccccccccccccccccccccccccc
cceeeeecccccccPFVPTMESFPPLDPFSPSLSSSPFFNNSNNTGMPFDPGFDLGLDNGLHSSCPPLFSQVQPGFEFGELEALcccccccccccccccccccccccchhhhhhhhhhhccc
ccccccccccccehhhcccchhhhhhhhhhTAAPGHVPFMEGPKGENPLERVQPTLYRKRRGTAEIPGLETVRRKGRKWEEGEAEEGSSADVDGSGLNYEccccchheeecccccccccccccccccccc
ccccccccccccchhhhhhhhccccccccccccccccccccccSDENDNGNNGLKRSENGEDHKGKRKGLPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIDcccccccceeeee
ccccttttccccccchhhhhhhhhhhhhhhhhhhhhhhhhcccccchhhhhhhhhhhhhYLKELLQRINDLHNELETTSAGSSLLPSASTSFQPLTPTLPTLPCRVKEELEKFTPTTPVSSIAVVYHSE
hhhhhhhhhhhhhhhhhhcccccccccccccccccccccccccccccccchhhhhcccccccccccccccccccSKKKPVEVRAREGRAVNIHMFCTRRPGLLLSTMRALDNLGLDVQQAVISCFNGFAL
DVFKAEQCREGQDVccccceeeeeecttceeeeeeeeccctthhhhhhhhhhhttccchhhhhhhhhttheeeeehhhccccccccLPEQIKAVLLDSAGFHVLLLLILLCRGLTVLAMIKchhhhhh
hhhhtttcchheehhhhhhttcceeeehh



8. VaICE1-HP indicates KOM38329.1 hypothetical protein LR48_Vigan03g171100 [Vigna angularis]

MLQQQPQQPQPQPQPPQLDSYFNFNINSAIPIPFVPTMESFPPLDPFSPSLSSSPFFNNSNNTGMPFDPGhccccccccccccccccchheeeccccccccccccccccccccccccccce
ecccccccccttFDLGLDNGLHSSCPPLFSQVQPGFEFGELEALTAAPGHVPFMEGPKGENPLERVQPTLYRKRRGTAEIPGccccccccccccccceecccccccccccchheccccchhhhc
ccccccccccccccchhhhhhhhhccccLETVRRKGRKWEEGEAEEGSSADVDGSGLNYESDENDNGNNGLKRSENGEDHKGKRKGLPAKNLMAERRRccccccccccccccccccccccccccc
ccccccccccccccccccccccccccccchhhhhhhhhhRKKLNDRLYMLRSVVPKISKMDRASILGDAIDYLKELLQRINDLHNELETTSAGSSLLPSASTSFQPLTPhhhhhhhhhhhhhh
hhhccccchhhhhhhhhhhhhhhhhhhhhhhhhcccccccccccccccccccccTLPTLPCRVKEELYPGTLPSPKNQPAKVEVRAREGRAVNIHMFCTRRPGLLLSTMRALDNLGLDVQQAVI
ccccccccccccccccccccccccccceeeeehcttceeeeeeeccccttthhhhhhhhhhhhtchhhhhhSCFNGFALDVFKAEQCREGQDVLPEQIKAVLLDSAGFHGMMhhhhhheeeeehhhc
ccccchhcchhhhhhhhhhhhccccccc

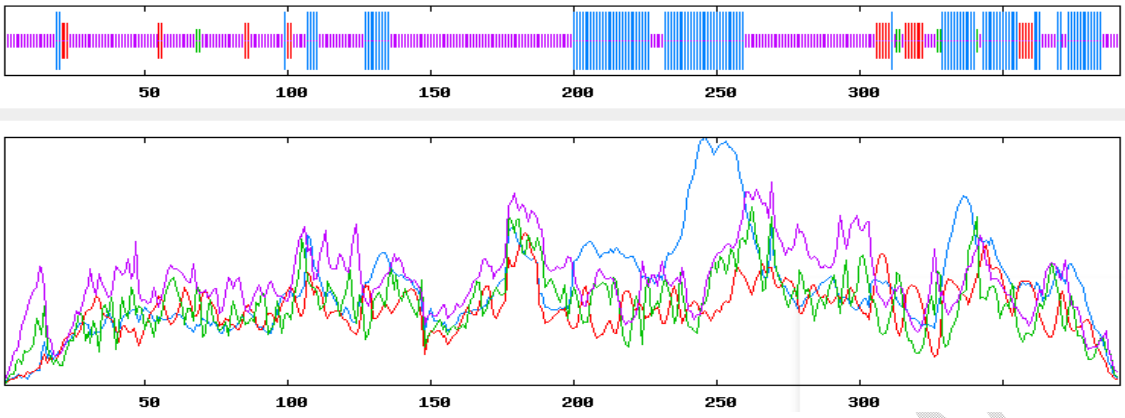
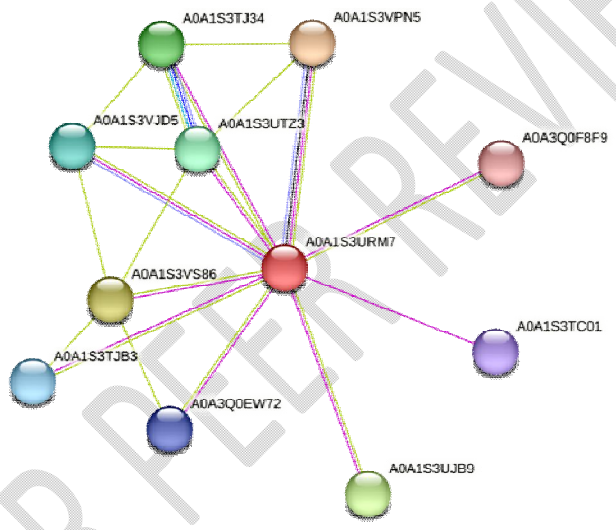


Fig 8. Protein-protein interaction network of ICE1 of *Vigna radiata.* Network nodes represent protein and edges represent protein-protein association.



Your Input:

A0A1S3URM7 *Transcription factor ICE1-like (467 aa)*

Predicted Functional Partners:

		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
A0A1S3VPN5	*Transcription factor FAMA isoform X1*					●		●	●	0.903
A0A1S3VS86	*E3 ubiquitin-protein ligase HOS1 isoform X1*					●		●		0.896
A0A1S3UJB9	*Heptahelical transmembrane protein 1*					●		●		0.885
A0A1S3TJ34	*Mitogen-activated protein kinase; Belongs to the protein kinase superfamily. Ser/Thr protein kinase family. MAP kinase s...*					●		●		0.883
A0A1S3UTZ3	*Mitogen-activated protein kinase; Belongs to the protein kinase superfamily. Ser/Thr protein kinase family. MAP kinase s...*					●		●		0.876
A0A1S3VJD5	*Transcription factor MUTE*					●		●	●	0.871
A0A1S3TJB3	*E3 SUMO-protein ligase SIZ1*					●		●		0.838
A0A3Q0EW72	*E3 SUMO-protein ligase SIZ1 isoform X1*					●		●		0.838
A0A1S3TC01	*Uncharacterized protein LOC106753921*					●				0.835
A0A3Q0F8F9	*Transcription factor bHLH95-like*					●		●		0.832

Your Current Organism:

Vigna radiata var. radiata

NCBI taxonomy Id: 3916

Other names: Phaseolus aureus, Phaseolus aureus Roxb., Phaseolus radiatus, Phaseolus radiatus L., V. radiata var. radiata, golden gram, green gram, mung bean

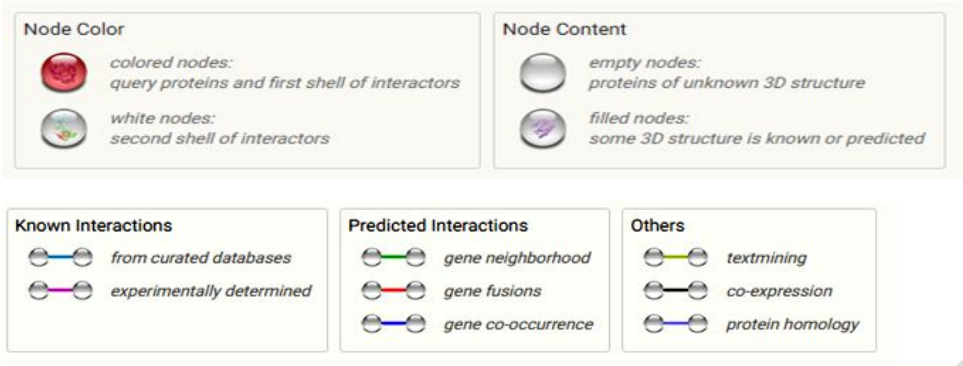


Fig. 9. Co-expression analysis of ICE1 protein along with are FAMA, HOS1, heptahelical transmembrane proteins, MAPKase, MUTE and E3 SUMO SIZ1 protein ligase in *Vigna radiata* and other legume sp.