Int.J.Curr.Microbiol.App.Sci (2019) *8*(7): 1279-1286

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 8 Number 07 (2019)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2019.807.152

Comprehensive Analysis of Universal Stress Proteins and their Promoter Sequences in Rice

Rakesh Bhowmick*

ICAR-Vivekananda Parvatiyakrishi Anusandhan Sansthan, Almora-263601, Uttarakhand, India
ICAR-Indian Agricultural Research Institute, New Delhi-110012, India

**Corresponding author*

Keywords

Universal stress proteins, Promoter sequences in rice

Article Info

Accepted:
12 June 2019
Available Online:
10 July 2019

A B S T R A C T

Rice is one of the important cereal crops and more than half of the world population is dependent on Rice as staple food. Rice cultivation is affected by various abiotic stresses like drought, salinity, cold etc. Universal Stress Proteins (*USP*s) are conserved proteins found across the plant, animals, fungi and bacteria. *USP*s are reported to be involved in various abiotic stress responses. In this study we have identified 43 *USP* genes in rice, distributed across the chromosome. Phylogenetic relationship between *OsUSP* protein sequences revealed four distinct groups, whereas phylogenetic tree of *OsUSP* promoter showed presence of six groups. Gene expression of *OsUSP*s in response to various abiotic stresses showed differential expression pattern in different stresses. We have also analysed promoter sequence of *OsUSP*s to find potential transcription factor binding site. We found 14533 TFBS in 43 *OsUSP* promoter region, which belong to 41 transcription factor family.

Introduction

Universal Stress Proteins (*USPs*) are highly conserved, found across most of the living organisms including bacteria, archea, fungi, plants and mammals (Kerk *et al.,* 2003; Ndimba *et al.,* 2005, Persson *et al.,* 2007). Initially the gene was first identified in a bacteria and it was named C13.5proteindue to its mobility in two dimensional gel electrophoresis. Later the name of these proteins was changed to Universal Stress Protein to reflect their responsiveness to cellular stress (Zarembinski *et al.,* 1998; Sousa and McKay, 2001). *USP*s are identified on the basis of presence of highly conserved 140-160 residues (PF00582). *E. coli* genome codes for six *USP*s and they are known to be involved in diverse biological processes including motility, adhesion, and resistance to oxidative stress (Nachin *et al.,* 2005). Orthologs of bacterial *USP*s are ubiquitously present in

plant genome. Arabidopsis and Barley genome is known to encode 44 and 66 *USP*s respectively (Kerk *et al.,* 2003; Li *et al.,* 2010).*USP* genes belong to two classes: ATP binding and ATP non-binding (Tkaczuk *et al.,* 2013). Though this gene family has been characterised in several plant species but function of *USP*s are still not completely understood. *USP*s play important role in various abiotic stresses. *SpUSP* was reported to reduce stomatal aperture to minimise harmful effect of drought stress (Rachid *et al.,* 2012). Arabidopsis *USP* gene, At3g53990 has chaperone function and expression is induced by heat and drought stress (Isokpehi *et al.,* 2012). Another ATP binding At*USP* gene, *At3g62550* is upregulated during drought stress. Arabidopsis *USP*s, AtPHOS32 and AtPHOS34 were found to be phosphorylated during microbial infection (Shinozaki and Yamaguchi-Shinozaki, 2007). Sauter *et al.,* (2002) demonstrated *OsUSP*1 is involved in activation of signalling cascade in response to ethylene during hypoxia. The *USP* genes of *Gossypium arboretum, Astragalus sinicus, Solanum pennellii*, and *Salicornia brachiate* were shown to be involved in water stress, nodulation, and drought, salt, and osmotic tolerances (Chou *et al.,* 2007; Maqbool *et al.,* 2009; Loukehaich *et al.,* 2012; Udawat *et al.,* 2014). Jung *et al.,* (2015) overexpressed At*USP* to show its role in heat and oxidative stress tolerance. In addition to stress response of *USP* genes, At*USP* promoter was shown to upregulate GUS expression during ABA, ACC, dehydration, heat, cold, salt, and osmotic stress (Bhuria *et al.,* 2016).

Rice is one of the major cereal crops cultivated throughout the world. Rice cultivation is affected by various biotic and abiotic stresses. No comprehensive study has been made till date to analyse promoter region of rice *USP* genes. In this study, we have identified 44 different *OsUSP* genes and we have identified common transcription factor binding site within promoters of *OsUSP*s. Gene expression of *OsUSP*s was analysed and distribution of *OsUSP*s across chromosomes was shown using mapchart software.

Materials and Methods

Identification of *OsUSP*s

For identification of candidate *USP* genes in rice, Basic Local Alignment Search Tool (BLAST) and Hidden Markov Model (HMM) search both was performed. Predicted protein sequences of rice were retrieved from ensemble plant database (https://plants.ensembl.org/index.html). Thirty-eight predicted Arabidopsis *USP* protein sequences were retrieved from TAIR database. Hidden Markov Model (HMM) profile of *USP* domain (PF00582) was fetched from PFam database (http://pfam.sanger.ac.uk/) and HMM search was employed against the local database of rice proteins using Hmmer 3.0 (e < 1e-5). As HMM profile search might fail to identify protein sequences containing incomplete *USP* domain, standalone BLASTP search was performed using Arabidopsis *USP* proteins as query against the respective rice proteins. E-value less than 1e-5 and percentage identity more than 50 was used as a threshold for further characterization. To confirm the results obtained using the HMMER algorithm and BLASTP search, the putative *USP* protein sequences were also queried against the Pfam and CDD (http://www.ncbi.nlm.nih.gov/cdd/) databases.

Phylogenetic analysis

The evolutionary history of *OsUSP* was inferred using the Neighbor-Joining method. The bootstrap consensus tree inferred from 1000 replicates is taken to represent the evolutionary history of the *OsUSP*. The evolutionary distances were computed using

the Poisson correction method and are in the units of the number of amino acid substitutions per site. The analysis involved 43 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 105 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 and Neighbor-Joining tree was reconstructed. Promoter sequences of *OsUSP*s were also retrieved and phylogenetic analysis of *OsUSP* promoters was also performed using neighbour-Joining method. Physical position of *OsUSP*s was fetched by BioMart data mining tool from Phytozome database. Chromosomal distribution of *OsUSP*s was visualised by MapChart software.

Promoter analysis of *OsUSP*s

For analysis of *OsUSP* gene promoters, 1Kb upstream sequence of all the *OsUSP*s were retrieved by using BioMart. For identification of transcription factor binding site within promoter, PlantPAN 2.0 software (http:// plantpan2.itps.ncku.edu.tw/) was used.

As *USP* genes are differentially expressed during abiotic stresses, we identified common transcription factor binding sites in promoters, using gene group analysis in PlantPAN 2.0 web tool.

Digital gene expression analysis

For gene expression analysis GENEVESTIGATOR software was used. Expression data of *OsUSP*s from Affymetrix rice genome array platform was fetched from GENEVESTIGATOR experiment dataset and analysed with Perturbation tool. Fold change of genes in response to various abiotic stresses was log transformed and expression change was represented in a heat map. Red and green colour represents up-regulation and down-regulation of genes respectively.

Results and Discussion

Identification of *OsUSP*s

HMMER and BLAST search was employed to identify putative *USP* genes in rice genome. Together HMMER and BLAST search identified 60 protein models containing *USP* domain. Based on locus information of the proteins, they were found to be coded by 43 loci. Thus, 43 non redundant *OsUSP*s were identified in rice genome, named *OsUSP*s1 to *OsUSP*43 according to chromosomal location in descending order (Table 1). Apart from *USP* domain, several *OsUSP* also contain kinase domain. *OsUSP*8, *OsUSP*9, *OsUSP*10, *OsUSP*15, *OsUSP*16, *OsUSP*21, *OsUSP*28, *OsUSP*29, *OsUSP*32, *OsUSP*33, *OsUSP*35, *OsUSP*36 and *OsUSP*38 have kinase domain in addition to *USP* domain, suggesting their putative role in signal transduction and activation of other protein function.

Phylogenetic analysis

The evolutionary history of both promoter sequence and *OsUSP* protein sequences were inferred by neighbour joining method (Fig. 1). After eliminating positions with gaps, a total of 650 positions were considered for calculation of evolutionary distance in promoter sequences. In case of protein sequences 105 amino acid positions were considered for distance matrix calculation. Phylogenetic analysis of *OsUSP* protein and promoter sequences revealed there is no correlation in evolutionary relationship between protein and promoter sequences. Promoter sequences were grouped into six distinct clades, whereas *OsUSP* protein sequences were grouped into four clades. Chromosomal distribution of *OsUSP*s revealed their non-uniform presence across the chromosomes (Fig. 2). There are 7 *OsUSP*s present in chromosome1 and 5; 8 *OsUSP*s on Chromosome 2; 6 in chromosome 3; 1 on

chromosome 9 and 11; 5 on chromosome 12. Chromosome 4 and 8 do not have any *OsUSP*. Tandem duplication of *OsUSP* genes were found on chromosome 2, chromosome 9 and chromosome 12.

Digital gene expression

Expression of *OsUSP*s in response to various abiotic stresses was analysed using GENEVESTIGATOR software (Fig. 3). We conducted differential expression of *OsUSP*s in response to anoxia, dehydration high temperature and cold. Expression pattern of *OsUSP*s differ in various stresses. *OsUSP*32 (LOC_Os07g47620) was found to be highly upregulated during all stresses examined except cold stress, whereas LOC_Os03g19270 (*OsUSP*17) was upregulated in dehydration and cold stress. LOC_Os02g47840 (*OsUSP12*) showed upregulation in response to dehydration stress but it was highly downregulated during anoxia, suggesting its differential mechanism of action in different stresses.

Expression of seven *OsUSP*s, LOC_Os02g 54590 (*OsUSP*15), LOC_Os01g39970 (*OsUSP*4), LOC_Os02g12660 (*OsUSP*9), LOC_Os03g40130 (*OsUSP*19), LOC_Os12g 07990 (*OsUSP*39), LOC_Os12g08060 (*OsUSP*40), LOC_Os10g32590 (*OsUSP*36) was not changed in response to any abiotic stresses, indicating their no role in stress response.*OsUSP*1 was found to be downregulated in all stresses except high temperature stress, where it is slightly upregulated.

Fig.1 Phylogenetic tree of *OsUSP* promoter (A) and protein (B) sequences. Neighbour-Joining method was employed for tree reconstruction

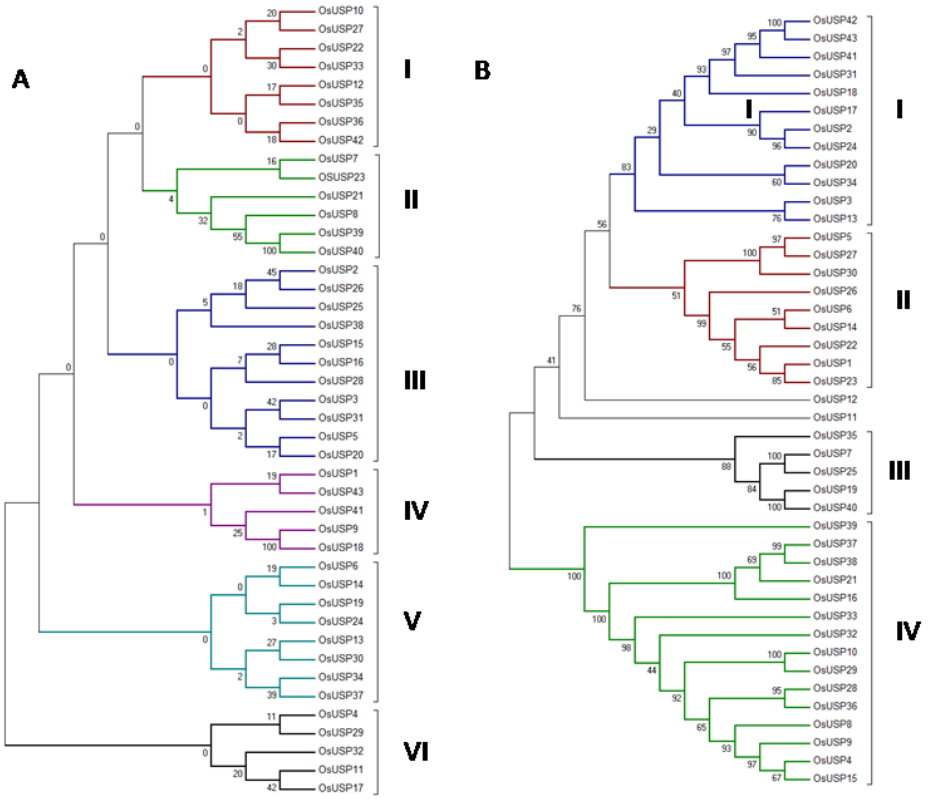

Fig.2 Chromosomal distribution of *OsUSP* genes. Left bar represent scale in Megabase (Mb)

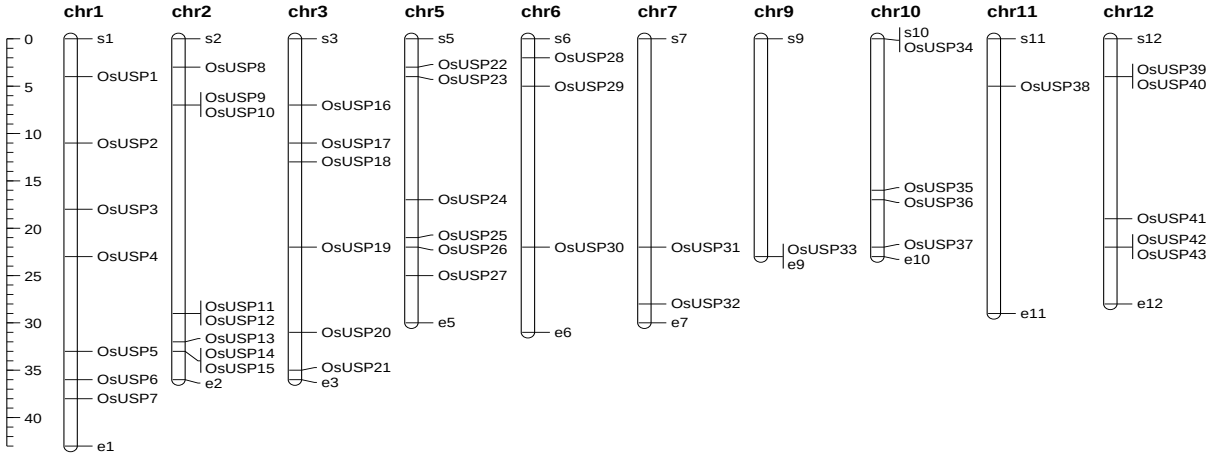

Fig.3 Expression analysis *OsUSP* genes in response to various abiotic stresses. Heat map of expression fold change was created using GENEVESTIGATOR software

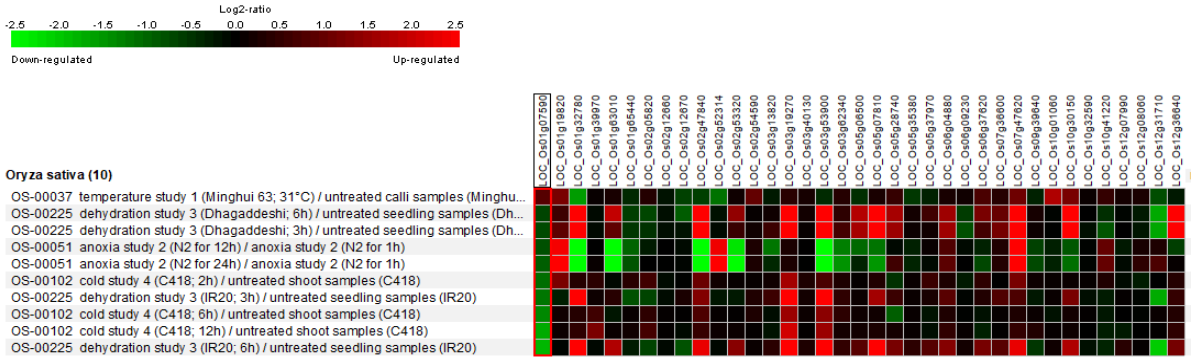

Fig.4 Number of different transcription factor binding site in OsUSP promoters. X axis represents number of TFBS

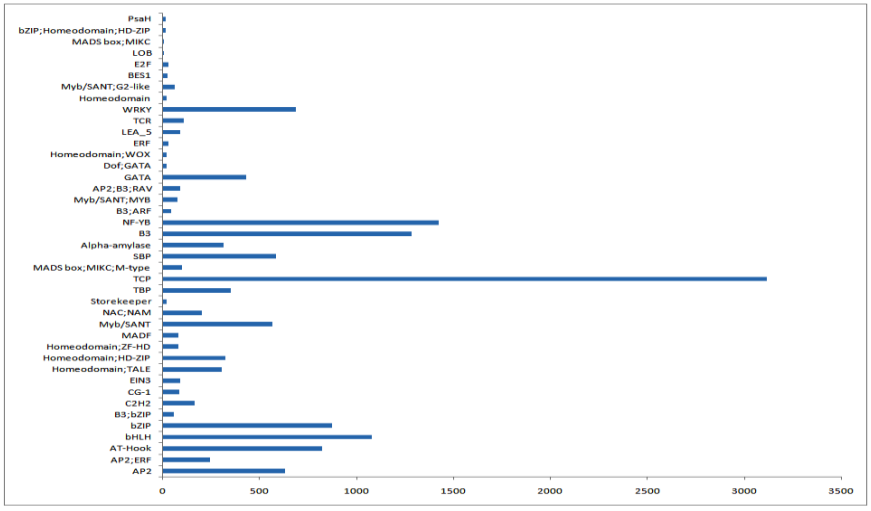

Table.1 Locus id of *OsUSP* genes

USP gene	Locus id	Chromosome	USP gene	Locus id	Chromosome
OsUSP1	LOC_Os01g07590	1	OsUSP23	LOC_Os05g07810	5
OsUSP2	LOC_Os01g19820	1	OsUSP24	LOC_Os05g28740	5
OsUSP3	LOC_Os01g32780	1	OsUSP25	LOC_Os05g35380	5
OsUSP4	LOC_Os01g39970	1	OsUSP26	LOC_Os05g37970	5
OsUSP5	LOC_Os01g57450	1	OsUSP27	LOC_Os05g42230	5
OsUSP6	LOC_Os01g63010	1	OsUSP28	LOC_Os06g04880	6
OsUSP7	LOC_Os01g65440	1	OsUSP29	LOC_Os06g09230	6
OsUSP8	LOC_Os02g05820	2	OsUSP30	LOC_Os06g37620	6
OsUSP9	LOC_Os02g12660	2	OsUSP31	LOC_Os07g36600	7
OsUSP10	LOC_Os02g12670	2	OsUSP32	LOC_Os07g47620	7
OsUSP11	LOC_Os02g47650	2	OsUSP33	LOC_Os09g39640	9
OsUSP12	LOC_Os02g47840	2	OsUSP34	LOC_Os10g01060	10
OsUSP13	LOC_Os02g52314	2	OsUSP35	LOC_Os10g30150	10
OsUSP14	LOC_Os02g53320	2	OsUSP36	LOC_Os10g32590	10
OsUSP15	LOC_Os02g54590	2	OsUSP37	LOC_Os10g41220	10
OsUSP16	LOC_Os03g13820	3	OsUSP38	LOC_Os11g08950	11
OsUSP17	LOC_Os03g19270	3	OsUSP39	LOC_Os12g07990	12
OsUSP18	LOC_Os03g22390	3	OsUSP40	LOC_Os12g08060	12
OsUSP19	LOC_Os03g40130	3	OsUSP41	LOC_Os12g31710	12
OsUSP20	LOC_Os03g53900	3	OsUSP42	LOC_Os12g36630	12
OsUSP21	LOC_Os03g62340	3	OsUSP43	LOC_Os12g36640	12
OsUSP22	LOC_Os05g06500	5			

Promoter analysis

Promoter region of gene is essential to drive expression of the genes. We found there is significant differential expression of *OsUSP* genes in response to various abiotic stresses as described above. In order to analyse promoter region of *OsUSP*s, we have retrieved 1Kb upstream sequence of respective *OsUSP* transcripts and this 1Kb region was considered as putative promoters of the genes. After scanning of promoter regions in PlantPAN database, we found 14533 transcription factor binding site across all the *OsUSP* promoter (Fig. 4). Forty one different transcription factors were found to bind with this TFBS. In our study we found highest number of TFBS for TCP (3118) transcription factor, whereas lowest number of TFBS was found for LOB and MIKC transcription factor. Both of these TFs have only six TFBS each in 43 *OsUSP* promoters. NF-YB, B-ZIP, B3, MYB, NAC, bHLH, AP2, WRKY and AT-Hook transcription factor binding sites were overrepresented in the *OsUSP* promoters. TCP, NF-YB and B3 TFBS were found to be present in all the *OsUSP*s.

In this present study 43 *OsUSP* genes were identified in rice genome and their promoter region was analysed to see which TFs bind with the promoters. Phylogenetic analysis of *USP* protein and promoter was done

separately and no correlation was found between NJ tree of *OsUSP* proteins and promoters. We found 14533 different transcription factors binding site in 43 *OsUSP*s, which belong to 41 TF family. Among these TFs, TCP, B-ZIP, WRKY, bHLH, GATA is already reported to be involved in various abiotic stress responses.

References

Bhuria, M., Goel, P., Kumar, S. and Singh, A.K., 2016. The promoter of AtUSP is co-regulated by phytohormones and abiotic stresses in Arabidopsis thaliana. *Frontiers in plant science*, *7*, p.1957.

Chou, M.X., Wei, X.Y., Chen, D.S. and Zhou, J.C., 2007. A novel nodule-enhanced gene encoding a putative universal stress protein from Astragalussinicus. *Journal of plant physiology*, *164*(6), pp.764-772.

Isokpehi, R.D., Simmons, S.S., Cohly, H.H., Ekunwe, S.I., Begonia, G.B. and Ayensu, W.K., 2011. Identification of drought-responsive universal stress proteins in viridiplantae. *Bioinformatics and biology insights*, *5*, pp.BBI-S6061.

Jung, Y.J., Melencion, S.M.B., Lee, E.S., Park, J.H., Alinapon, C.V., Oh, H.T., Yun, D.J., Chi, Y.H. and Lee, S.Y., 2015. Universal stress protein exhibits a redox-dependent chaperone function in Arabidopsis and enhances plant tolerance to heat shock and oxidative stress. *Frontiers in plant science*, *6*, p.1141.

Kerk, D., Bulgrien, J., Smith, D.W. and Gribskov, M., 2003. Arabidopsis proteins containing similarity to the universal stress protein domain of bacteria. *Plant Physiology*, *131*(3), pp. 1209-1219.

Li, W.T., Wei, Y.M., Wang, J.R., Liu, C.J., Lan, X.J., Jiang, Q.T., Pu, Z.E. and Zheng, Y.L., 2010. Identification, localization, and characterization of putative USP genes in barley. *Theoretical and Applied Genetics*, *121*(5), pp.907-917.

Loukehaich, R., Wang, T., Ouyang, B., Ziaf, K., Li, H., Zhang, J., Lu, Y. and Ye, Z., 2012. SpUSP, an annexin-interacting universal stress protein, enhances drought tolerance in tomato. *Journal of experimental botany*, *63*(15), pp.5593-5606.

Loukehaich, R., Wang, T., Ouyang, B., Ziaf, K., Li, H., Zhang, J., Lu, Y. and Ye, Z., 2012. SpUSP, an annexin-interacting universal stress protein, enhances drought tolerance in tomato. *Journal of experimental botany*, *63*(15), pp.5593-5606.

Maqbool, A., Zahur, M., Husnain, T. and Riazuddin, S., 2009. GUSP1 and GUSP2, two drought-responsive genes in Gossypiumarboreum have homology to universal stress proteins. *Plant molecular biology reporter*, *27*(1), pp.109-114

Nachin, L., Nannmark, U. and Nyström, T., 2005. Differential roles of the universal stress proteins of Escherichia coli in oxidative stress resistance, adhesion, and motility. *Journal of bacteriology*, *187*(18), pp.6265-6272.

Ndimba, B.K., Chivasa, S., Simon, W.J. and Slabas, A.R., 2005. Identification of Arabidopsis salt and osmotic stress responsive proteins using two- dimensional difference gel electrophoresis and mass spectrometry. *Proteomics*, *5*(16), pp. 4185-4196.

Persson, Ö., Valadi, Å., Nyström, T. and Farewell, A., 2007. Metabolic control of the Escherichia coli universal stress protein response through fructose- 6- phosphate. *Molecular microbiology*, *65*(4), pp.968-978.

Sauter, M., Rzewuski, G., Marwedel, T. and Lorbiecke, R., 2002. The novel

ethylene- regulated gene OsUsp1 from rice encodes a member of a plant protein family related to prokaryotic universal stress proteins. *Journal of Experimental Botany*, *53*(379), pp.2325-2331.

Shinozaki, K. and Yamaguchi-Shinozaki, K., 2007. Gene networks involved in drought stress response and tolerance. *Journal of experimental botany*, *58*(2), pp.221-227.

Sousa, M.C. and McKay, D.B., 2001. Structure of the universal stress protein of *Haemophilus influenzae*. *Structure*, *9*(12), pp.1135-1141.

Tkaczuk, K.L., A. Shumilin, I., Chruszcz, M., Evdokimova, E., Savchenko, A. and Minor, W., 2013. Structural and functional insight into the universal stress protein family. *Evolutionary applications*, *6*(3), pp.434-449.

Udawat, P., Mishra, A. and Jha, B., 2014. Heterologous expression of an uncharacterized universal stress protein gene (SbUSP) from the extreme halophyte, *Salicornia brachiata*, which confers salt and osmotic tolerance to E. coli. *Gene*, *536*(1), pp.163-170.

Zarembinski, T.I., Hung, L.W., Mueller-Dieckmann, H.J., Kim, K.K., Yokota, H., Kim, R. and Kim, S.H., 1998. Structure-based assignment of the biochemical function of a hypothetical protein: a test case of structural genomics. *Proceedings of the National Academy of Sciences*, *95*(26), pp.15189-15193.

How to cite this article:

Rakesh Bhowmick. 2019. Comprehensive Analysis of Universal Stress Proteins and Their Promoter Sequences in Rice. *Int.J.Curr.Microbiol.App.Sci.* 8(07): 1279-1286.
doi: https://doi.org/10.20546/ijcmas.2019.807.152