
RESEARCH ARTICLE

***In silico* analysis of ERF family in Arabidopsis, Oryza and Solanum crop plants**

Shouhartha Choudhury

Department of Biotechnology, Assam University, Silchar-788011, Assam, India

Corresponding authors email Id: shouharthac@gmail.com

Manuscript received: March 11, 2019; Decision on manuscript: April 30, 2019; Manuscript accepted: July. 9, 2019

Abstract

In silico efforts were made to get an insight of ERF family in *Arabidopsis* a model plant, *Oryza* and *Solanum* cultivable crop plants. The present investigation identified 179, 188, 171 AP2 domains in *Arabidopsis thaliana*, *Oryza sativa* and, *Solanum lycopersicum* respectively. The expression study demonstrated that ERF109 (query gene) is highly expressed in the flowering stage and lower in the germinating stage of the plants and found to be abundant in flower, blade, hypocotyl, lateral root, and cotyledon. The principle of transcription factor data analysis suggested that multiple hits of 7 ERF109 (RRTF1) genes and 127, 140, 136 AP2 domains in *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* respectively. Several abiotic stress responsive genes of the ERF family have been identified. These results will be useful future functional analyses of the ERF family in plants. Also the results may be useful for the biotechnologist and plant breeders in understanding the insight of ERF family in *Arabidopsis* and for the genetic improvement in *Oryza* and *Solanum* crops.

Key words: *In silico*, ERF family, Arabidopsis, Oryza, Solanum

Introduction

The abiotic stress conditions such as freezing, drought, salt, low oxygen, oxidative, osmotic, heat, ABA ethylene, jasmonic acid, and abscisic acids have adverse effects on the growth and production of plants. To survive, grow and reproduce under environments that fluctuate annually and diurnally of plants have evolved complicated systems that regulate adaptation in response to stress. The aspects of adaptation processes, such as developmental, physiological, and biochemical changes are regulated by the stress-responsive gene. The transcription factors play central roles in gene expression by regulating downstream genes as trans-acting elements via specific binding to cis-acting elements in the promoters of target genes. The identification of stress-responsive promoters, cis-acting and trans-acting elements involved in transcriptional responses of stress-responsive genes have been identified (Yamaguchi-Shinozaki and Shinozaki 2006). The APETALA 2/ethylene-responsive element binding factor (AP2/ERF) family is a large group of plant-specific transcription factors that includes four major subfamilies: the AP2, RAV, and ERF with special regulation of dehydration-responsive element-binding (DREB) subfamilies (Sakuma, Liu *et. al.*, 2002). The ERF subfamily members binding to the ethylene-responsive element (ERE) and the regulation of DREB

subfamily members have been established that the major factors involved in plant abiotic stress responses by regulating via the cis-acting dehydration-responsive element/C-repeat (DRE/CRT) element (Sakuma, Maruyama *et al.*, 2006). The ERF family contains a single AP2 domain and classified into two major groups ERF and CBF/DREB (Sakuma, Liu *et al.*, 2002, Dubouzet, Sakuma *et al.*, 2003, Sakuma, Maruyama *et al.*, 2006). The single AP2 domain first identified a conserved DNA-binding domain namely ethylene-responsive element binding 1, 2, 3, and 4 specifically bind to the GCC box (Ohme-Takagi and Shinshi 1995, Fujimoto, Ohta *et al.*, 2000). The sequence contains 268 amino acids with 64 are binding to the specific DNA called AP2 domain. The *Arabidopsis thaliana* genome sequencing was completed in 2000 (Arabidopsis Genome Initiative, 2000) Total 145 genes postulated containing AP2/ERF family with 83% (121 genes) belonging to ERF family (Initiative 2000, Dubouzet, Sakuma *et al.*, 2003). Most of the members of the ERF family have yet to be studied, despite the likelihood that these genes play many important roles in physiological aspects in plants. A great deal of transgenic experiment will be required to determine the specific biological function of each gene in the ERF family. The phylogeny analysis has become apparent that a large gene family of transcription factors consists of subfamily those are closely related to each other (Kranz, Denekamp *et al.*, 1998, Bailey, Martin *et al.*, 2003, Reyes, Muro-Pastor *et al.*, 2004). Comparative and functional analysis of each transcription factor belonging to the ERF family should be done, taking into account for functional redundancy. This process and assessment of relationships between gene families would provide a guide for predicting the functions of a specific gene induces in the genome. The current availability of *Oryza sativa* and *Solanum lycopersicum* genome sequences allowed comparative and functional analysis

between *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum*, which is useful in terms of functional and evolutionary diversity of transcription factor family. In this study, an establishment and overall information of the ERF family in *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* was attempted. The ERF family in *Arabidopsis thaliana* genome was surveyed again, based on the observation ERF family and special regulation of the DREB subfamily presented. A comparative and functional analysis between *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* was performed. The overall results represent a function and regulation of ERF and DREB subfamily members in response to abiotic stress in plants.

Materials and Methods

Sequence and database

The genome sequence of *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* were downloaded from TAIR and Ensemble genome database. *The primary sequence of Arabidopsis was downloaded from the TAIR database (Garcia-Hernandez, Berardini et al., 2002). SMART (Letunic, Doerks et al., 2011) and Pfam (Bateman, Coin et al., 2004) were retrieved to identify the AP2 domain. PROSITE (Hulo, Bairoch et al. 2006) performed for identification of the domain, family, and functional sites as well as associated pattern and profile. PROCHECK (Laskowski, MacArthur et al., 2006) examine the stereochemical quality of the primary peptide sequence.*

Standalone tools and GO annotation

HMMER (Zhang and Wood 2003) executed using multiple sequence alignments of the specific domain as a profile search. HMMER is a statistical algorithm, making multiple sequence alignment (MSA) of the specific domain as a profile search, its implements methods using probabilistic models called the profile hidden

Markov model. Standalone BLAST performed for homologs gene in all organisms. BLAST2GO (Conesa and Götz 2008) performed for the sequences accuracy of specific transcription factor in the genome. BLAST2GO is bioinformatics and computational tool for high-throughput gene annotation of the novel sequence data. The functional information retrieves via Gene Ontology (GO) annotation, a controlled vocabulary of the functional attributes.

Domain, motif, phylogeny and gene expression

Multiple sequence alignment (MSA) methods to calculate the best match of the homologs sequences. MSA of multiple hits sequence analysis was carried out with a web-based tool MultAlin (Corpet 1988) for identification of conserved domain. MEGA7 (Kumar, Stecher *et. al.*, 2016) performed for the phylogenetic tree using Neighbor-Joining Methods. MEME (Bailey, Boden *et. al.*, 2009) suite is a computational tool for discovery and analysis of sequence motifs; a web-based tool for retrieving specific motifs in given sequences. Expression analysis carried out using GENEVESTIGATOR (Zimmermann, Hirsch-Hoffmann *et. al.*, 2004) tools is a high-performance search engine for gene expression across different biological contexts.

Results and discussion

In the present study the genome sequences of *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* were downloaded and performed the gene ontology annotation along

with identification of domain, motifs, and phylogeny and gene expression studies. MSA of multiple hits sequence analysis was carried out for identification of conserved domain. Transcription factor Id: AT4G344110.1 (Common name: ERF109, F10M10, and RRTF1; Organisms: *Arabidopsis thaliana*, Family: ERF; Description: Redox responsive transcription factor 1) was identified. The primary nucleotide and peptide sequence of ERF109 retrieved from the TAIR database. The coding sequence composed of 1386 nucleotides and 268 peptides with 64 initiated DNA binding domains, which are well known as an AP2 domain (Table1). Genome sequence of *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* downloaded from different specialization database and executed HMMER algorithm obtain total of 179, 188, 171 AP2 domain in *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* respectively. Standalone BLAST results represent 127, 129, 125 homologs gene in *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* respectively. Gene ontology annotation results demonstrated the sequence accuracy of the ERF109 genes in *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* (Table2). Multiple sequence alignment result demonstrated conserved domain, and high consensus indicates extended AP2 domain (Fig.1) and their specific motifs (Fig.2). The phylogenetic tree demonstrated the evolutionary relationship between ERF genes in *Solanum lycopersicum*. Particular clades define multi-functional genes in the ERF family response to abiotic stress in plants (Fig.3).

Table 1: Gene and query sequence

Gene Id	829591
TF Id	AT4G344110.1
Common name	ERF109, F10M10, and RRTF1
Organisms	<i>Arabidopsis thaliana</i>
Family	ERF
Description	Redox responsive transcription factor 1

Table2: Gene ontology annotation summary of the *Solanum lycopersicum*

Gene Id	Gene	Protein
Solyc10g050970.1.1	ERF109	ethylene-responsive transcription factor ERF109
Solyc01g108240.1.1	ERF109	ethylene-responsive transcription factor ERF109
Solyc02g090790.1.1	ERF109	ethylene-responsive transcription factor ERF109
Solyc02g090800.1.1	ERF109	ethylene-responsive transcription factor ERF109
Solyc02g090770.1.1	ERF071	ethylene-responsive transcription factor ERF071
Solyc09g075420.1.1	ERF071	ethylene-responsive transcription factor ERF071
Solyc06g068830.1.1	ERF113	ethylene-responsive transcription factor ERF113
Solyc03g118190.1.1	ERF113	ethylene-responsive transcription factor ERF113
Solyc12g056590.1.1	ABR1	ethylene-responsive transcription factor ABR1
Solyc04g071770.1.1	ABR1	ethylene-responsive transcription factor ABR1
Solyc04g051360.1.1	ABR1	ethylene-responsive transcription factor ABR1
Solyc04g012050.1.1	ABR1	ethylene-responsive transcription factor ABR1
Solyc12g042210.1.1	ERF114	ethylene-responsive transcription factor ERF114
Solyc07g053740.1.1	ERF004	ethylene-responsive transcription factor 4
Solyc10g009110.1.1	ERF004	ethylene-responsive transcription factor 4
Solyc03g006320.1.1	ERF004	ethylene-responsive transcription factor 4
Solyc02g077810.1.1	ERF004	ethylene-responsive transcription factor 4
Solyc03g093540.1.1	ERF005	ethylene-responsive transcription factor 5
Solyc03g093560.1.1	ERF005	ethylene-responsive transcription factor 5
Solyc05g052040.1.1	ERF005	ethylene-responsive transcription factor 5
Solyc08g078190.1.1	ERF005	ethylene-responsive transcription factor 5
Solyc03g093550.1.1	ERF005	ethylene-responsive transcription factor 5
Solyc01g065980.1.1	RAP2-3	ethylene-responsive transcription factor RAP2-3
Solyc10g006130.1.1	ERF003	ethylene-responsive transcription factor 3
Solyc07g064890.1.1	ERF003	ethylene-responsive transcription factor 3
Solyc07g049490.1.1	ERF003	ethylene-responsive transcription factor 3
Solyc12g005960.1.1	ERF003	ethylene-responsive transcription factor 3
Solyc12g038450.1.1	ERF003	ethylene-responsive transcription factor 3
Solyc06g065820.1.1	ERF003	ethylene-responsive transcription factor ERF003
Solyc06g068360.1.1	ERF003	ethylene-responsive transcription factor ERF003
Solyc03g117130.1.1	ERF003	ethylene-responsive transcription factor ERF003
Solyc12g038440.1.1	ERF003	ethylene-responsive transcription factor 3
Solyc08g078170.1.1	ERF002	ethylene-responsive transcription factor 2
Solyc08g078180.1.1	ERF002	ethylene-responsive transcription factor 2
Solyc03g093610.1.1	ERF002	ethylene-responsive transcription factor 2
Solyc01g090320.1.1	ERF002	ethylene-responsive transcription factor 2
Solyc01g090340.1.1	ERF002	ethylene-responsive transcription factor 2
Solyc08g007230.1.1	ERF002	ethylene-responsive transcription factor 2
Solyc03g123500.1.1	RAP2-12	ethylene-responsive transcription factor RAP2-12
Solyc06g063070.1.1	RAP2-12	ethylene-responsive transcription factor RAP2-12
Solyc12g049560.1.1	RAP2-12	ethylene-responsive transcription factor RAP2-12
Solyc05g052050.1.1	ERF001	ethylene-responsive transcription factor 1
Solyc04g014530.1.1	ERF001	ethylene-responsive transcription factor 1
Solyc05g051200.1.1	ERF001	ethylene-responsive transcription factor 1
Solyc11g011740.1.1	ERF001	ethylene-responsive transcription factor 1
Solyc09g066360.1.1	ERF001	ethylene-responsive transcription factor 1
Solyc09g089930.1.1	ERF001	ethylene-responsive transcription factor 1
Solyc11g006050.1.1	ERF001	ethylene-responsive transcription factor 1
Solyc11g011750.1.1	ERF001	ethylene-responsive transcription factor 1
Solyc03g005520.1.1	ERF001	ethylene-responsive transcription factor 1
Solyc01g090310.1.1	ERF013	ethylene-responsive transcription factor 13
Solyc01g090300.1.1	ERF013	ethylene-responsive transcription factor 13
Solyc01g090370.1.1	ERF013	ethylene-responsive transcription factor 13
Solyc05g050790.1.1	ERF013	ethylene-responsive transcription factor 13
Solyc02g077370.1.1	PTI5	pathogenesis-related genes transcriptional activator
Solyc05g052030.1.1	ERF106	ethylene-responsive transcription factor ERF106
Solyc03g093530.1.1	ERF106	ethylene-responsive transcription factor ERF106

Solyc06g054630.1.1	ERF017	ethylene-responsive transcription factor ERF017
Solyc12g009240.1.1	ERF017	ethylene-responsive transcription factor ERF017
Solyc06g051840.1.1	CRF4	ethylene-responsive transcription factor CRF4
Solyc03g007460.1.1	CRF4	ethylene-responsive transcription factor CRF4
Solyc02g077360.1.1	ERF091	ethylene-responsive transcription factor ERF091
Solyc09g066350.1.1	ERF098	ethylene-responsive transcription factor ERF098
Solyc09g089920.1.1	ERF098	ethylene-responsive transcription factor ERF098
Solyc03g005510.1.1	ERF098	ethylene-responsive transcription factor ERF098
Solyc06g082590.1.1	PTI6	pathogenesis-related genes transcriptional activator PTI6
Solyc01g095500.1.1	PTI6	pathogenesis-related genes transcriptional activator PTI6
Solyc11g042580.1.1	ERF021	ethylene-responsive transcription factor ERF021
Solyc08g081960.1.1	CRF2	ethylene-responsive transcription factor CRF2
Solyc08g078420.1.1	ERF038	ethylene-responsive transcription factor ERF038
Solyc08g078410.1.1	ERF038	ethylene-responsive transcription factor ERF038
Solyc12g056980.1.1	RAP2-4	ethylene-responsive transcription factor RAP2-4
Solyc04g072900.1.1	RAP2-4	ethylene-responsive transcription factor RAP2-4
Solyc06g035700.1.1	ERF025	ethylene-responsive transcription factor ERF025
Solyc04g054910.1.1	RAP2-13	ethylene-responsive transcription factor RAP2-13
Solyc09g091950.1.1	ERF062	ethylene-responsive transcription factor ERF062
Solyc03g005500.1.1	ERF096	ethylene-responsive transcription factor ERF096
Solyc09g066340.1.1	ERF096	ethylene-responsive transcription factor ERF096
Solyc02g093130.1.1	ERF011	ethylene-responsive transcription factor ERF011
Solyc12g013660.1.1	ERF053	ethylene-responsive transcription factor ERF053
Solyc07g054220.1.1	ERF054	ethylene-responsive transcription factor ERF054
Solyc04g078640.1.1	ERF010	ethylene-responsive transcription factor ERF010
Solyc01g091760.1.1	ERF010	ethylene-responsive transcription factor ERF010
Solyc06g050520.1.1	DREB	dehydration responsive element binding
Solyc07g042230.1.1	ERF016	ethylene-responsive transcription factor ERF016
Solyc01g067540.1.1	ERF086	ethylene-responsive transcription factor ERF086
Solyc05g050830.1.1	ERF039	ethylene-responsive transcription factor ERF039
Solyc01g090560.1.1	ERF039	ethylene-responsive transcription factor ERF039
Solyc04g080910.1.1	DREB2D	dehydration-responsive element-binding 2D
Solyc05g052410.1.1	DREB2A	dehydration-responsive element-binding 2A
Solyc04g050750.1.1	DREB2G	dehydration-responsive element-binding 2G
Solyc10g080310.1.1	DREB2F	dehydration-responsive element-binding 2F
Solyc02g067020.1.1	DREB2C	dehydration-responsive element-binding 2C
Solyc10g076370.1.1	DREB2A	dehydration-responsive element-binding 2A
Solyc10g076380.1.1	DREB2C	dehydration-responsive element-binding 2C
Solyc09g009240.1.1	ERF024	ethylene-responsive transcription factor erf024
Solyc12g008350.1.1	TINY	ethylene-responsive transcription factor TINY
Solyc03g120840.1.1	TINY	ethylene-responsive transcription factor TINY
Solyc12g044390.1.1	DREB3	dehydration-responsive element-binding 3
Solyc08g066660.1.1	DREB3	dehydration-responsive element-binding 3
Solyc06g066540.1.1	DREB3	dehydration-responsive element-binding 3
Solyc04g072300.1.1	LEP	ethylene-responsive transcription factor LEP
Solyc11g045680.1.1	LEP	ethylene-responsive transcription factor LEP
Solyc11g045690.1.1	LEP	ethylene-responsive transcription factor LEP
Solyc12g009490.1.1	WIN1	ethylene-responsive transcription factor WIN1
Solyc03g116610.1.1	WIN1	ethylene-responsive transcription factor WIN1
Solyc03g119580.1.1	CRF6	ethylene-responsive transcription factor CRF6
Solyc03g026270.1.1	DREB1A	dehydration-responsive element-binding 1A
Solyc01g009440.1.1	DREB1F	dehydration-responsive element-binding 1F
Solyc12g056430.1.1	DREB1E	dehydration-responsive element-binding 1E
Solyc08g007820.1.1	DREB1E	dehydration-responsive element-binding 1E
Solyc08g007830.1.1	DREB1E	dehydration-responsive element-binding 1E
Solyc11g012980.1.1	ERF014	ethylene-responsive transcription factor ERF014
Solyc11g042560.1.1	ERF022	ethylene-responsive transcription factor ERF022
Solyc05g009450.1.1	ERF118	ethylene-responsive transcription factor ERF118
Solyc04g007180.1.1	ERF118	ethylene-responsive transcription factor ERF118

Solyc10g080650.1.1	RAP2-11	ethylene-responsive transcription factor RAP2-11
Solyc10g078610.1.1	RAP2-11	ethylene-responsive transcription factor RAP2-11
Solyc01g005630.1.1	RAP2-11	ethylene-responsive transcription factor RAP2-11
Solyc03g114440.1.1	RAP2-11	ethylene-responsive transcription factor RAP2-11
Solyc08g082210.1.1	ERF061	ethylene-responsive transcription factor ERF061
Solyc10g083560.1.1	ERF024	ethylene-responsive transcription factor ERF024
Solyc03g117230.1.1	ERF084	ethylene-responsive transcription factor ERF084
Solyc08g080290.1.1	ERF027	ethylene-responsive transcription factor ERF027
Solyc03g119800.1.1	ERF043	ethylene-responsive transcription factor ERF043
Solyc01g014720.1.1	POS1	physalis organ size 1
Solyc01g008880.1.1	POS1	physalis organ size 1
Solyc05g013540.1.1	ESR2	ethylene-responsive transcription factor ESR2
Solyc01g057080.1.1	ERF020	ethylene-responsive transcription factor ERF020
Solyc02g030210.1.1	AP2-like	AP2-like ethylene-responsive transcription factor

Fig.1: Extended AP2 domain

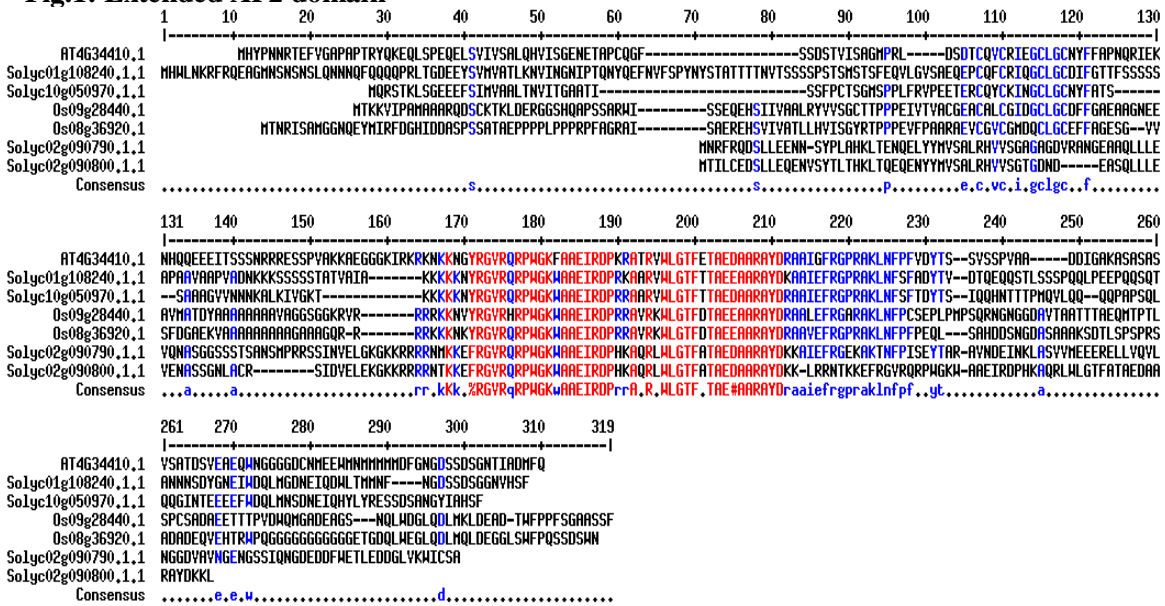
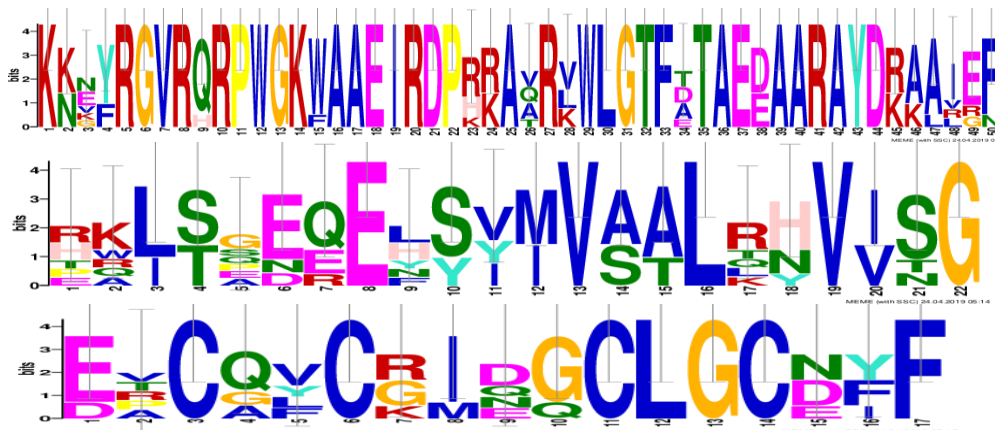


Fig2: Specific motifs of AP2 domain



The expression study demonstrated ERF109 (query gene) is highly expressed in the flowering stage and lower in the germinating stage of the plants and found to be abundant in flower, blade, hypocotyl, lateral root, and cotyledon. The principle of transcription factor data analysis suggested multiple hits of 7 ERF109 (RRTF1) genes and 127, 140, 136 AP2 domains in *Arabidopsis thaliana*, *Oryza sativa*, and *Solanum lycopersicum* respectively (Table3).

ERF family is a responsible to the abiotic stress and based on the expression data it may be suggested that query gene in ERF family is highly express in plants. A closer look at ERF

family and analyzed genes known so far, those have a different composition of the functional domain. The present investigation summarized 179, 188, 171 AP2 domains in *Arabidopsis thaliana*, *Oryza sativa* and, *Solanum lycopersicum* respectively. Some other specific domain also involved, which cannot be consider as an ERF family gene. Removed all other specific domain involved genes and listed only single AP2 domain involved genes. The AP2 domain is a central domain in AP2/ERF family. The AP2/ERF family is a large transcription factor family contains subfamily AP2, ERF, RAV, and DREB.

Fig 3: ERF family in *Solanum lycopersicum*

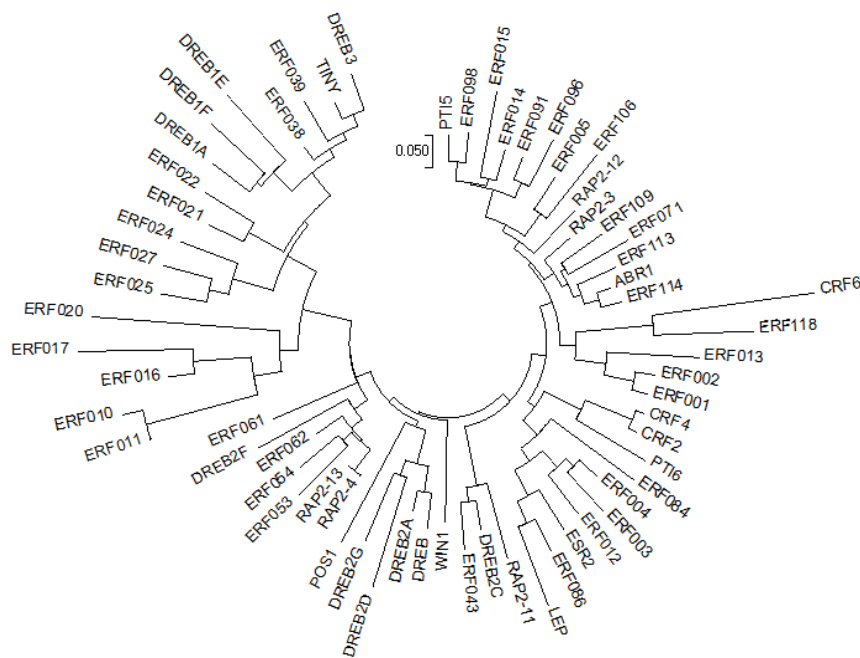


Table 3: Summary of the AP2 domain

Organisms	HMMER Hits	BLAST Hits
<i>Arabidopsis thaliana</i>	179	127
<i>Oryza sativa</i>	188	129
<i>Solanum lycopersicum</i>	171	125
Total	538	381

Table 4: The function of ERF transcription factor family

Species	Gene	Functions	Reference
<i>Arabidopsis thaliana</i>	TINY	Negatively regulates development, positively regulates abiotic stress	Sun <i>et. al.</i> , 2008
<i>Arabidopsis thaliana</i>	RAP2.4	Positively regulates light, ethylene-mediated development, drought stress response	Lin <i>et. al.</i> , 2008
<i>Arabidopsis thaliana</i>	ERF4	Positively regulates ethylene and ABA responses	Yang <i>et. al.</i> , 2009
<i>Arabidopsis thaliana</i>	RAP2.3	Positively regulates low oxygen, oxidative and osmotic stress responses	Yang <i>et. al.</i> , 2009
<i>Arabidopsis thaliana</i>	ERF53	Positively regulates heat and ABA response	Gong <i>et. al.</i> , 2008
<i>Arabidopsis thaliana</i>	ERF1	Positively regulates salt, drought, and heat response	Yang <i>et. al.</i> , 2009
<i>Arabidopsis thaliana</i>	DREB2A	Positively regulates drought, salt, heat, cold response	Sakuma <i>et. al.</i> , 2002
<i>Oryza sativa</i>	DREBs	Positively regulates drought, salt, heat, cold response	Dubouzet <i>et. al.</i> , 2003
<i>Solanum lycopersicum</i>	ERF3	Positively regulates drought, salt, cold, and osmotic stress response	Wu <i>et. al.</i> , 2008
<i>Solanum lycopersicum</i>	ERF2	Positively regulates freezing and ethylene response	Hongxing <i>et. al.</i> , 2005

The present study identified abiotic stress responsive genes related to ERF family with special regulation of DREB subfamily in *Arabidopsis thaliana*, *Oryza sativa* and, *Solanum lycopersicum*. Identified genes regulate the abiotic stress response such as freezing, drought, salt, low oxygen, oxidative, osmotic, heat, ABA ethylene, jasmonic acid, and abscisic acids (Table 5 -7). The highest hits obtained in *Oryza sativa* and lowest in *Solanum lycopersicum*. The previously identified gene in *Oryza sativa*: Os02g42580, Os02g52880, Os03g02650, Os04g36640, Os04g48330, Os06g42910 and Os12g07030 (Fig. 5) remark, there is no AP2 domain which is characteristic of the AP2/ERF family. *Oryza sativa*: OsERF#010 (Os06g09690) and OsERF#139 (Os06g09730) was completely identical; afterward they provided a gene Id: Os06g09717, I identified that gene in LOC_Os06g09717.1. The previous reports suggested, Sakuma *et. al.*,

reported 17, 121 and 6 genes in AP2, ERF, RAV family in *Arabidopsis thaliana*, and Nakano *et. al.*, reported 18, 122, 6 in AP2, ERF, and RAV family in *Arabidopsis thaliana* and compared with 139 ERF genes in *Oryza sativa* (Riechmann, Heard *et. al.*, 2000, Sakuma, Liu *et. al.*, 2002, Nakano, Suzuki *et. al.*, 2006, Sakuma, Maruyama *et. al.*, 2006). Therefore family-wise classification study is an essential component for a better understanding of specific transcription factor family in the genome. In conclusion, efforts were made to get an insight of ERF family in *Arabidopsis*, *Oryza* and *Solanum* crop plants. Several abiotic stress responsive genes of ERF family have been identified. These findings will be useful for the biotechnologist and plant breeders for understanding the insight of the ERF family in *Arabidopsis* and for the genetic improvement in *Oryza* and *Solanum* crops.

Table 5: Biological function of ERF family transcription factor

Cis-element	Plant species	Protein	References
Drought Responsive element (DRE/CRT) 5'TACCGACAT3'	<i>Arabidopsis thaliana</i>	DREBs CBFs	(Stockinger, Gilmour <i>et. al.</i> , 1997, Jaglo-Ottosen, Gilmour <i>et. al.</i> , 1998, Haake, Cook <i>et. al.</i> , 2002)
Drought Responsive element (DRE/CRT) 5'TACCGACAT3'	<i>Oryza sativa</i>	DREBs	(Dubouzet, Sakuma <i>et. al.</i> , 2003)
Ethylene Responsive Element (ERE or GCC box) 5'TAAGAGCCGCC3' or 5'AGCCGCC3'	<i>Arabidopsis thaliana</i>	ERF1-5 EBP	(Büttner and Singh 1997, Fujimoto, Ohta <i>et. al.</i> , 2000)

Table 6: Cis-acting elements and their interacting motifs

Plant species	Gene name	Interacting motifs	Regulated traits	References
<i>Arabidopsis thaliana</i>	TINY	DRE/CRT & GCC	Negatively regulates development, positively regulates abiotic stress	Sun <i>et. al.</i> , 2008
<i>Arabidopsis thaliana</i>	RAP2.4	DRE/CRT & GCC	Positively regulates light and ethylene mediated development and drought stress response	Lin <i>et. al.</i> , 2008
<i>Arabidopsis thaliana</i>	ERF4	DRE/CRT & GCC	Positively regulates ethylene and ABA responses	Yang <i>et. al.</i> , 2009
<i>Arabidopsis thaliana</i>	RAP2.3	DRE/CRT & GCC	Positively regulates low oxygen, oxidative, and osmotic stress responses	Yang <i>et. al.</i> , 2009
<i>Arabidopsis thaliana</i>	ERF53	DRE/CRT & GCC	Positively regulates heat and ABA response	Gong <i>et. al.</i> , 2008
<i>Arabidopsis thaliana</i>	ERF1	DRE/CRT & GCC	Positively regulates salt, drought, and heat response	Yang <i>et. al.</i> , 2009
<i>Arabidopsis thaliana</i>	DREB2A	DRE/CRT & GCC	Positively regulates drought, salt, heat, cold response	Sakuma <i>et. al.</i> , 2002
<i>Solanum lycopersicum</i>	ERF3	DRE/CRT & GCC	Positively regulates drought, salt, cold, and osmotic stress response	Wu <i>et. al.</i> , 2008

Table7: Function of ERF transcription factor family

Gene	Functions	Species
ABR1	Abscisic acid response, sugar signaling	<i>Arabidopsis thaliana</i>
RAP2.3	Positively regulates low oxygen, oxidative, and osmotic stress responses	<i>Arabidopsis thaliana</i>
ERF53	Positively regulates heat and ABA response	<i>Arabidopsis thaliana</i>
ERF1	Positively regulates salt, drought, and heat response	<i>Arabidopsis thaliana</i>
DREB2A	Positively regulates drought, salt, heat, cold response	<i>Arabidopsis thaliana</i>
RAP2.4	Positively regulates light- and ethylene-mediated Growth regulation	<i>Arabidopsis thaliana</i>
TINY	Growth regulation	<i>Arabidopsis thaliana</i>
ERF4	Ethylene, jasmonic acid, and abscisic acid response	<i>Arabidopsis thaliana</i>
ERF7	Abscisic acid response	<i>Arabidopsis thaliana</i>
WIN1	Wax accumulation	<i>Arabidopsis thaliana</i>

Competing financial interest's statement

The author did not avail of any financial assistance from any source in undertaking the present study. The author states that there is no dispute of interests concerning the publication of this paper.

References

- Bailey, P. C., Martin C, Toledo-Ortiz, G., Quail, P. H., Huq E., Heim M. A., Jakoby M., Werber, M. and Weisshaar B. 2003. Update on the basic helix-loop-helix transcription factor gene family in *Arabidopsis thaliana*. *The Plant Cell*, 15(11): 2497-2502.
- Bailey, T. L., Boden, M., Buske, F. A., Frith, M. Grant, C. E., Clementi, L., Ren, J., Li, W. W. and Noble, W. S. 2009. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, 37:(suppl_2): 202-208.

Acknowledgments

The author is thankful to Assam University, Silchar, Assam, India for providing the requisite lab facilities in carrying out this research work and also thankful to the editor and anonymous reviewer for critical comments and suggestions.

- Bateman, A., Coin, L., Durbin R., Finn, R. D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S. and Sonnhammer, E. L. 2004. The Pfam protein families database. *Nucleic Acids Res.*, 32(suppl_1): 138-141.
- Büttner, M. and Singh, K. B. 1997. *Arabidopsis thaliana* ethylene-responsive element binding protein (AtEBP), an ethylene-inducible, GCC box DNA-binding protein interacts with an Ocs element binding protein. *Proceed. National Academy Sci.*, 94(11): 5961-5966.
- Conesa, A. and Götz, S. 2008. Blast2GO: A comprehensive suite for functional analysis

- in plant genomics. *Int. J. Plant Genomics*, 619832.
6. Corpet, F. 1988. Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.*, 16(22): 10881-10890.
 7. Dubouzet, J. G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E. G., Miura, S., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. 2003. OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought, high-salt and cold-responsive gene expression. *The Plant J.*, 33(4): 751-763.
 8. Fujimoto, S. Y., Ohta, M. A., Usui, Shinshi, H. and Ohme-Takagi, M. 2000. Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. *The Plant Cell*, 12(3): 393-404.
 9. Garcia-Hernandez, M., Berardini, T., Chen, G., Crist, D., Doyle, A., Huala, E., Knee, E., Lambrecht, M., Miller, N. and Mueller, L. A. 2002. TAIR: a resource for integrated Arabidopsis data. *Funct. Integrative Genomics*, 2(6): 239-253.
 10. Haake, V., D. Cook, J. Riechmann, O. Pineda, M. F. Thomashow and J. Z. Zhang. 2002. Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis. *Plant Physio.*, 130(2): 639-648.
 11. Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., Castro, E. De, Langendijk-Genevaux, P. S., Pagni, M. and Sigrist, C. J. 2006. The PROSITE database. *Nucleic Acids Res.*, 34(suppl_1): 227-230.
 12. Jaglo-Ottosen, K. R., Gilmour, S. J., Zarka, D. G., Schabenberger, O., and Thomashow, M.F. 1998. Arabidopsis CBF1 over expression induces COR genes and enhances freezing tolerance. *Science*, 280(5360): 104-106.
 13. Kranz, H. D., Denekamp, M., Greco, R., Jin, H., Leyva, A., Meissner, R. C., Petroni, K., Urzainqui, A. Bevan, M. and Martin, C. 1998. Towards functional characterisation of the members of the R2R3-MYB gene family from Arabidopsis thaliana. *The Plant J.*, 16(2): 263-276.
 14. Kumar, S., Stecher, G. and Tamura, K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol., Bio. Evol.*, 33(7): 1870-1874.
 15. Laskowski, R.A, MacArthur, M.W., Thornton J.M. 2001. PROCHECK: validation of protein structure coordinates. In: Rossmann MG, Arnold E (eds) *International tables of crystallography, volume F. Crystallography of biological macromolecules.* Kluwer Academic Publishers, Dordrecht, pp 722–725.
 16. Letunic, I., Doerks, T. and Bork, P.. 2011. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res.*, 40(D1): D302-D305.
 17. Nakano, T., Suzuki, K., Fujimura, T. and Shinshi, H. 2006. Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physio.*, 140(2): 411-432.
 18. Ohme-Takagi, M. and Shinshi, H. 1995. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *The Plant Cell*, 7(2): 173-182.
 19. Reyes, J. C., Muro-Pastor, M. I. and Florencio, F. J. 2004. The GATA family of transcription factors in Arabidopsis and rice. *Plant Physio.*, 134(4): 1718-1732.
 20. Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C. Z., Keddie, J. Adam, L., Pineda, O., Ratcliffe, O. and Samaha, R. 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science*, 290(5499): 2105-2110.
 21. Sakuma, Y., Liu, Q., Dubouzet, J. G., Abe, H., Shinozaki, K. and Yamaguchi-Shinozaki, K. 2002. DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration and cold-inducible gene

- expression. *Bioch. Biophysical Res. Communi.*, 290(3): 998-1009.
22. Sakuma, Y., Maruyama, K. Osakabe, Y., Qin, F., Seki, M., Shinozaki, K. and Yamaguchi-Shinozaki, K. 2006. Functional analysis of an *Arabidopsis* transcription factor, DREB2A, involved in drought-responsive gene expression. *The Plant Cell*, 18(5): 1292-1309.
 23. Stockinger, E. J., Gilmour, S. J. and Thomashow, M. F. 1997. *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proceed. National Academy Sci.*, 94(3): 1035-1040.
 24. The *Arabidopsis* Genome Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408: 796–815
 25. 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.
 26. Zhang, Z. and Wood, W. I. 2003. A profile hidden Markov model for signal peptides generated by HMMER. *Bioinformatics*, 19(2): 307-308.
 27. Zimmermann, P., Hirsch-Hoffmann, M. Hennig, L. and Grissem, W. 2004. GENEVESTIGATOR *Arabidopsis* microarray database and analysis toolbox. *Plant Physio.*, 136(1): 2621-2632.