GENOME–WIDE ANALYSIS OF ETHYLENE RESPONSIVE FACTOR IN MAIZE: AN IN SILICO APPROACH

HUSSAIN, A.[1] – SHAHID, M.[1] – MUSTAFA, G.[1] – AZEEM, M.[3] – SHAHEEN, H. L.[3*] – SULTAN, S.[1] – HUSSAIN, S.[1] – AHMED, I.[2]

[1]*Department of Bioinformatics and Biotechnology, Government College University, Faisalabad 38000, Pakistan*

[2]*School of Chemistry and Environment, Beihang University, Beijing 100191, China*

[3]*Department of Botany, Government College University, Faisalabad 38000, Pakistan*

**Corresponding author*
e-mail: huma.muzammal@yahoo.com
tel: +92-336-7735200

(Received 27th Apr 2016; accepted 24th Aug 2016)

Abstract. Transcription factors are usually considered as key player for gene regulation. Among various transcription factor families, AP2/ERF superfamily is well known for regulating various stress responses in plants. The family encompasses AP2/ERF domain, which is involved in DNA binding comprises of about 60 to 70 amino acids. To date, there is no detailed report presenting structural and functional prediction of ERF genes in *Zea mays* (L.). The current study presents a comprehensive genome-wide analysis of 105 ERF genes in maize (ZmERF) using several computational techniques. We performed phylogenetic analysis, conserved motif analysis, chromosomal localization, gene structure analysis and multiple sequence alignment of ERF genes. The phylogenetic analysis led to classification of these ERF members into 10 major groups and various subgroups and inferred evolutionary relationship among the groups on the basis of various protein motifs as well as intron/exon structure. The mapping of ERF genes on 10 maize chromosomes revealed their existence on all chromosomes with most number (17 genes) carried by chromosome 1 and least number (8 genes) found on chromosome 3. Interestingly, a very limited intron frequency was resulted in gene structure analysis. Gene ontology analysis concludes that ZmERF are involved in responses to various stresses including both biotic and abiotic. The results of the present study provide important structural information to design functional analyses of the ERF genes in *Z. mays.*

Keywords: *in silico study, phylogeny, gene structure analysis, chromosomal mapping, protein structure, motif analysis, Z. mays*

Introduction

Abiotic stresses are serious threat to agricultural crops and reduce their yield up to 50 % (Boyer, 1982). These stresses are regulated at cellular, biochemical, physiological and molecular level by adapted plants. Various stress responsive genes associated with molecular signaling have been discovered which are regulated and expressed at molecular level (Seki et al., 2003; Zhang et al., 2004). Transcription factors are referred as DNA binding proteins, which bind at cis-regulatory element and initiate transcription process. On the basis of function difference and similarities, transcription factors are classified into many families AP2/ERF, WRKY, ARF, FAR1 etc (Pabo and Sauer, 1992).

AP2/ERF play a key role in gene expression, related to cell proliferation, hormone secretion, plant reproduction, and biotic as well as abiotic stress responses (Hattori et al., 2009; Hinz et al., 2010). In past decade, AP2/ERF family has become more attentional gene family. AP2/ERF superfamily originated as a result of horizontal

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

transfer from bacteria/viruses to plants (Magnani, 2004; Shigyo et al., 2006). This is one of the most important transcription factors having AP2-DNA binding domain (Kim et al., 2011), which consists of 60 to 70 amino acid residues first discovered in *Arabidopsis thaliana* (Jofuku et al., 1994). Later on, these genes were cloned in tobacco plants (Takagi and Shinshi, 1995). At present, AP2/ERF gene have been identified in several species of plants including barley, soya bean, grape, poplar, wheat, foxtail millet, peach, sorghum, brassica, maize, moso bamboo (Zhang et al., 2008; Zhang et al., 2012; Song et al., 2013; Yan et al., 2013; Du et al., 2014; Lata et al., 2014; Wu et al., 2015). AP2/ERF superfamily is further divided into three families: AP2, ERF, and RAV (http://planttfdb.cbi.pku.edu.cn) on the basis of the domain observed in it. AP2 family have two AP2/ERF domain and possess very important function in regulation of developmental process like leaf formation, flower (Elliott et al., 1996), embryo, ovule (Boutilier et al., 2002; Song et al., 2013) and fruit development (Zhang et al., 2012). AP2 family is further classified into AP2 and ANT subfamilies (Shigyo et al., 2006). ERF family have single AP2/ERF domain and carry crucial function in ethylene signal transduction in response to environmental stresses (Takagi and Shinshi, 1995; Dubouzet et al., 2003), pathogen related stimuli and regulated pathogenesis-related gene expression (Zarei et al., 2011). Recently, overexpression of ERF gene was studied in rice (Zhang and Huang, 2010), tomato and tobacco under saline and drought environment (Guo et al., 2004; Zhang and Huang, 2010). In RAV, in addition to one AP2/ERF domain, there is another DNA binding B3-like domain, which is plant specific binding and also present in other transcription factors (Kagaya, et al., 1999). Moreover, RAV also show ethylene response as well as brassinosteroid hormonal response (Alonso, 2003; Hu et al., 2004).

Further studies on ERF family shows its further division into two subfamilies: the CBF/DREB (dehydration response element binding) and ERF (Ethylene responsive factors) (Sakuma et al., 2002). These two sub families are further arranged into many groups, 15 in Rice and 12 in *Arabidopsis* (Nakano et al., 2006), 10 in grape (Licausi et al., 2010) 10 in cucumber (Hu and Liu, 2011) and 107 in *Z. mays* (Huang et al., 2014). Even high conservation in domain sequence, each family shows different DNA binding site. ERF subfamilies bind to the GCC box (AGCCGCC) (Takagi et al., 1995), where G2, G5 and C7 are considered to be the important residues (Hao et al., 2002), DREB bind TACCGACAT where C4, G5, and C7 are basic residues for binding (Jiang et al., 1996; Sakuma et al., 2002). AP2 binds to GCAC(A/G)N(A/T)TCCC(A/G)ANG(C/T) (Gong et al., 2008; Wilson and Krizek, 2000). The ERF-associated amphiphilic repression (EAR) motif was also identified in many species of plant like *Arabidopsis*, *Z. mays* and sorghum (Yan et al., 2013). The three dimensional analysis of AP2/ERF protein domain shows it consist of 3 anti-parallel beta-sheet and one alpha-helix (Yamasaki et al., 2013) arginine and tryptophan in beta-sheet have basic function in formation of GCC box binding domain and serine/threonine also important residue for DNA binding domain (Shanker et al., 2012).

Though, AP2/ERF superfamily has been thoroughly studied, yet there is no detailed report in *Z. mays* which could provide information about the structure, characteristic and function of each gene keeping in view the contribution of ERF members in plant stress regulation. Current study is focused on genome-wide analysis of ERF subfamily in *Z. mays*, which is the most important cereal crops in the world after wheat and rice. The crop has a nutritional value vital in our daily life (Verheye, 2010). The study encompasses computational analysis of ZmERF *viz.* analysis of conserved domain,

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

phylogenetic evolutionary studies, gene structure analysis, chromosomal distribution, gene ontology and BLAST hits distribution of ERF genes among different plant species. This prediction-based study might help to design wet-lab experiments against abiotic stress-resistance in *Z. mays*. The study also extends significant information related to ZmERF genes to make possible solution of constraints in growing the crop in unfavorable ecological conditions.

Materials and methods

Identification and bioinformatics analysis of ZmERF protein

AP2/ERF superfamily is already studied in *Z. mays* which identified 107 ERF subfamily members (Du et al., 2014). These 107 ERF genes accession were used for current study. Sequences and related data of these genes were retrieved from different databases including Plant Transcription Database (http://planttfdb.cbi.pku.edu.cn/) (Jin et al., 2014), Phytozome (http://phytozome.jgi.doe.gov/) and maize GDB (http://www.maizegdb.org/) (Andorf et al., 2015). As result we were able to collect data and sequences of 105 ERF genes, excluding two ERF genes; GRMZM2G061227 and GRMZM2G16097I which were not found from these databases. New Annotation was also given according to their chromosomal location e.g GRMZM1.8 and GRMZM 6.5 is given to GRMZM2G039112 and GRMZM2G085964, respectively, where digits before point represents chromosome number and after point location of gene on respective chromosome in ascending order. Bioinformatics analysis of ERF genes was performed which involve amino acid (a.a) length, molecular weight (kDa), isoelectric point (Ip) using ExPASy server (http://www.expasy.ch/tools).

Conserved domain analysis and phylogenetic analysis

To observe conserved residues in AP2 domain, multiple sequence alignment of these 105 ERF genes was executed using CLC work bench software package (Knudsen et al., 2011) with parameters: gap opening cost: 10.0, gap extension cost: 0.1, end gap cost: free. The aligned sequences of ERF genes were used to construct phylogenetic tree in CLC viewer 7.6 (Knudsen et al., 2011) with neighbor joining method, Distance measure: Jukes-Cantor and Bootstrap: 1000 Replicates.

Analysis of ZmERF proteins motifs and gene ontology annotation

Motif analysis was conducted using MEME online software (Bailey and Elkan, 1994; Bailey et al., 2006). Attributes used to accomplish data analysis include; number of different motifs: 20, minimum motif width: 15, maximum motif width: 54 amino acids, distribution of the motif occurrence: zero or one per sequences. The gene ontology of ERF protein sequence was performed using BLAST GO (Conesa and Götz, 2008; Gotz et al., 2011) with default parameters. GO analysis of ZmERF described the biological process, species distribution, molecular function and cellular localization (Gotz et al., 2011). Databases employed to search out the sequence homologies were: NCBI non-redundant protein (Nr), NCBI nucleotide sequence (Nt), Protein family, Kyoto Encyclopedia of Genes and Genomes (KEGG), Swiss-PROT protein, Cluster of Orthologous Groups (COGs) and Gene Ontology (GO).

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

Chromosomal distribution, gene structure and sub-cellular localization of ZmERF genes

Location of all these genes on chromosomes is determined as per data of Phytozome and NCBI for instance, chromosome number, chromosome length and start position of gene. Graphical map is made by designing scale on the longest chromosome basis (chromosome number 5), and gene position on chromosome is considered as the gene annotation: like GRMZM1.14, GRMZM1.4 and GRMZM6.5 for AC206031, AC206951 and GRMZM2G085964, respectively. The genomic data was retrieved from phytozome and NCBI for gene gene structure analysis. Longest gene was found after sequence analysis to design a scale. The graphical representation of exon and intron was constructed according to numeric value and position of exon taken from NCBI. Prediction of sub-cellular localization was carried by two online bioinformatics tools namely WoLF PROST and Plant-mPLoc (Horton et al., 2007; Chou and Shen, 2010).

Results

Bioinformatics analysis of ethylene responsive factors

Comparative analysis of ERF genes in different species are given in *Table 1*, showing *Z. mays*, sorghum and peach have almost similar number of ERF gene e.g 107, 105 and 104 respectively. Chines cabbage and carrot have higher number of ERF, while *Arabidopsis*, moso bamboo, and foxtail millet have fewer number than *Z. mays* (*Table 1*). The individual genes of *Z. mays* are listed in *Table 2*, with their predicted features, including annotation chromosome number, exon number, intron number, gene length (ORF), amino acid (a.a) sequence length, number of exons, and isoelectric point. The length of genes ranged from 392 bp (GRMZM1.1) to 6700 bp (GMZM1.6), amino acid sequence length 130 (GRMZM1.1) to 2000 (GMZM1.6) a.a, exon number 1 to 6 (GRMZM1.7) and Ip ranged from 4.23 (GRMZM2.5) to 13.03 (GRMZM6.5) (*Table 2*).

Table 1. *Comparison of AP2/ERF genes in different species of plants*

Plants	***Total AP2/ERF***	***Classification***					***Reference***
		AP2	***ERF***	***DREB***	***RAV***	***Soloist***	
Foxtail millet	171	28	90	48	5	*	(Lata et al. 2014)
Chines cabbage	291	49	139	109	14	1	(Song, Li, and Hou 2013)
Peach	131	21	104	*	5	1	(C. H. Zhang et al. 2012)
Sorghum	126	16	105	*	4	1	(Yan et al. 2013)
Z. mays	184	22	107	51	3	1	(Du et al. 2014)
Moso bamboo	116	28	80	*	7	1	(Wu et al. 2015)
Carrot	267	38	143	71	12	3	(Yao et al. 2015)
Arabidopsis	147	18	65	57	6	1	(Sakuma et al. 2002)
Vitis venefera	149	20	86	36	6	1	(Licausi et al. 2010)
Oryza sativa	164	26	79	52	7	-*	(Wu et al. 2015)
Musa acuminata	265	67	119	81	16	3	(Lakhwani et al. 2016)
Musa balbisiana	318	71	144	99	22	4	(Lakhwani et al. 2016)

*Data not found in cited paper

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

Table 2. *Ethylene responsive factors and their predicted characteristic in Z. mays*

New annotation	Accession #	ORF * Length	a.a*	Ip*	Chr #	Exons #	Intron #
GRMZM1.1	AC198979	392	130	8.46	1	1	0
GRMZM1.2	GRMZM2G129674	1667	302	5.15	1	3	2
GRMZM1.3	GRMZM2G018984	2001	340	4.91	1	2	1
GRMZM1.4	AC206951	482	160	11.65	1	1	0
GRMZM1.5	AC200038	1152	339	7.18	1	2	1
GRMZM1.6	GRMZM2G703514	6700	2000	10.58	1	5	4
GRMZM1.7	GRMZM2G461905	3971	1185	10.31	1	6	5
GRMZM1.8	GRMZM2G039112	1836	414	4.37	1	1	0
GRMZM1.9	GRMZM2G016079	837	195	10.94	1	1	0
GRMZM1.10	GRMZM2G480434	608	202	4.83	1	1	0
GRMZM1.11	GRMZM2G033656	857	242	7.93	1	1	0
GRMZM1.12	GRMZM2G309731	510	160	10.88	1	1	0
GRMZM1.13	GRMZM2G369472	1137	269	6.37	1	1	0
GRMZM1.14	AC206031	839	170	10.82	1	1	0
GRMZM1.15	GRMZM2G010100	1256	418	4.77	1	1	0
GRMZM1.16	GRMZM2G009598	1320	238	11.43	1	1	0
GRMZM1.17	GRMZM2G381441	1183	285	9.95	1	1	0
GRMZM2.1	GRMZM2G067463	623	177	5.35	2	2	1
GRMZM2.2	GRMZM2G068967	756	222	10.23	2	1	0
GRMZM2.3	GRMZM2G475678	1070	356	4.58	2	1	0
GRMZM2.4	GRMZM2G055180	1336	279	5.18	2	1	0
GRMZM2.5	GRMZM2G079825	1365	271	4.23	2	1	0
GRMZM2.6	GRMZM2G087059	2385	445	9.94	2	2	1
GRMZM2.7	GRMZM2G148333	2313	423	4.57	2	2	1
GRMZM2.8	GRMZM2G138396	1074	263	6.28	2	1	0
GRMZM2.9	GRMZM2G011236	1219	311	8.12	2	1	0
GRMZM2.10	GRMZM2G125460	1138	284	9.05	2	1	0
GRMZM2.11	GRMZM2G458437	939	296	7.06	2	1	0
GRMZM3.1	GRMZM2G142179	1519	354	4.49	3	1	0
GRMZM3.2	GRMZM2G133168	4945	201	12.14	3	1	1
GRMZM3.3	GRMZM2G169382	1274	227	7.34	3	2	1
GRMZM3.4	GRMZM2G105266	1839	299	7.58	3	1	0
GRMZM3.5	GRMZM2G149756	3202	357	6.24	3	2	1
GRMZM3.6	GRMZM2G002119	849	186	10.73	3	1	0
GRMZM3.7	GRMZM2G310368	1047	237	9.02	3	1	0
GRMZM3.8	GRMZM2G474326	1148	229	9.12	3	1	0
GRMZM4.1	AC213666	491	163	7.64	4	1	0
GRMZM4.2	GRMZM2G089995	955	203	7.35	4	1	0
GRMZM4.3	GRMZM2G060206	625	124	9.09	4	1	0
GRMZM4.4	GRMZM2G165257	1295	294	8.30	4	1	0
GRMZM4.5	GRMZM2G018398	2689	361	4.44	4	2	2
GRMZM4.6	GRMZM2G173771	1052	230	8.50	4	1	0
GRMZM4.7	GRMZM2G060465	823	253	7.87	4	1	0
GRMZM4.8	GRMZM2G076896	2637	329	730	4	2	1
GRMZM4.9	GRMZM2G146028	1284	208	8.14	4	1	0
GRMZM4.10	GRMZM2G052720	1413	196	5.29	4	1	0
GRMZM5.1	GRMZM2G024871	522	137	7.53	5	1	0
GRMZM5.2	GRMZM2G322672	2499	393	9.48	5	4	3
GRMZM5.3	GRMZM2G079653	1109	211	7.88	5	1	0
GRMZM5.4	GRMZM2G085678	1672	213	10.03	5	2	1
GRMZM5.5	GRMZM2G073258	1003	290	6.40	5	1	0
GRMZM5.6	GRMZM2G016434	2720	408	8.86	5	2	1

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

GRMZM5.7	GRMZM2G047999	941	301	4.40	5	1	0
GRMZM5.8	GRMZM2G466044	1529	294	5.81	5	1	0
GRMZM5.9	GRMZM2G057386	1525	329	6.10	5	1	0
GRMZM5.10	GRMZM2G103085	1219	236	8.94	5	1	0
GRMZM5.11	GRMZM2G110333	2746	363	4.69	5	2	2
GRMZM6.1	GRMZM2G135452	989	173	6.96	6	2	1
GRMZM6.2	GRMZM2G328197	969	316	4.50	6	1	0
GRMZM6.3	GRMZM2G175543	1698	197	6.35	6	1	0
GRMZM6.4	GRMZM2G106591	1406	230	10.47	6	2	1
GRMZM6.5	GRMZM2G085964	1225	223	13.03	6	2	1
GRMZM6.6	GRMZM2G100727	1348	297	8.57	6	1	0
GRMZM6.7	GRMZM2G020016	1815	210	10.92	6	1	0
GRMZM6.8	GRMZM2G020150	1857	220	8.20	6	1	0
GRMZM6.9	GRMZM2G317596	1008	290	6.94	6	1	0
GRMZM7.1	GRMZM2G379652	989	329	6.34	7	1	0
GRMZM7.2	GRMZM2G478965	1088	362	5.79	7	1	0
GRMZM7.3	GRMZM2G019443	1022	322	9.64	7	1	0
GRMZM7.4	GRMZM2G171569	1449	307	10.06	7	1	1
GRMZM7.5	AC233933	788	262	10.25	7	1	0
GRMZM7.6	GRMZM2G123119	998	257	6.51	7	1	0
GRMZM7.7	GRMZM2G060517	608	136	8.45	7	1	0
GRMZM7.8	GRMZM2G363052	3043	283	4.42	7	1	4
GRMZM7.9	GRMZM2G052667	3074	418	5.17	7	2	2
GRMZM7.10	GRMZM2G131281	1041	227	9.79	7	2	1
GRMZM7.11	GRMZM2G018138	737	218	10.42	7	1	0
GRMZM7.12	GRMZM2G384386	1425	308	7.18	7	1	0
GRMZM7.13	GRMZM2G307119	1496	315	7.11	7	1	0
GRMZM7.14	GRMZM2G025062	1211	235	6.29	7	2	1
GRMZM8.1	GRMZM2G081892	2010	336	4.75	8	1	1
GRMZM8.2	AC187157	1101	291	7.68	8	2	1
GRMZM8.3	GRMZM2G044077	1597	325	4.43	8	1	0
GRMZM8.4	GRMZM2G053503	1177	230	8.87	8	2	1
GRMZM8.5	GRMZM2G021790	854	281	7.59	8	1	0
GRMZM8.6	GRMZM2G120401	1295	294	8.30	8	1	0
GRMZM8.7	GRMZM2G457562	1657	235	6.97	8	1	0
GRMZM8.8	GRMZM2G066158	1235	205	10.69	8	1	0
GRMZM8.9	GRMZM2G174347	1192	233	8.46	8	1	0
GRMZM8.10	GRMZM2G151542	2021	303	4.41	8	1	0
GRMZM8.11	GRMZM2G132223	855	165	10.29	8	1	0
GRMZM8.12	GRMZM2G132185	1121	241	9.36	8	1	0
GRMZM9.1	GRMZM2G171179	2960	363	4.97	9	2	2
GRMZM9.2	GRMZM2G156006	762	169	6.26	9	2	1
GRMZM9.3	GRMZM2G429378	947	315	4.75	9	1	0
GRMZM10.1	GRMZM2G544539	2253	349	9.42	10	2	1
GRMZM10.2	GRMZM2G020054	1295	186	7.92	10	1	0
GRMZM10.3	GRMZM2G055070	593	186	7.20	10	1	0
GRMZM10.4	GRMZM2G425798	2212	422	10.10	10	3	2
GRMZM10.5	GRMZM2G023708	1029	272	4.35	10	1	0
GRMZM10.6	GRMZM2G175525	761	253	9.88	10	1	0
GRMZM10.7	GRMZM2G080516	1955	270	5.73	10	1	0
GRMZM10.8	GRMZM2G438202	1671	222	10.52	10	1	0
GRMZM10.9	GRMZM2G104260	904	172	6.67	10	2	1
GRMZM10.10	GRMZM2G164591	1115	178	10.30	10	1	0

**ORF=Origin of replication frame, A.A= Amino acid, Ip=Isoelectric Point , Chr=Chromosome*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

Multiple sequence alignment

Analysis of AP2/ERF conserved domain showed that most of the sequences have conserved amino acid residues. The motif AAEIR is almost conserved in all amino acid sequences of ZmERF. Consensus sequence established that the active site Y41 is highly conserved in all sequences. G4, A37 and A38 sites also harbor conserved residues in most of the ERF protein sequences. At the N-terminal, residues are more conserved as compared to C-terminal residues of the domain (*Fig. 1*).

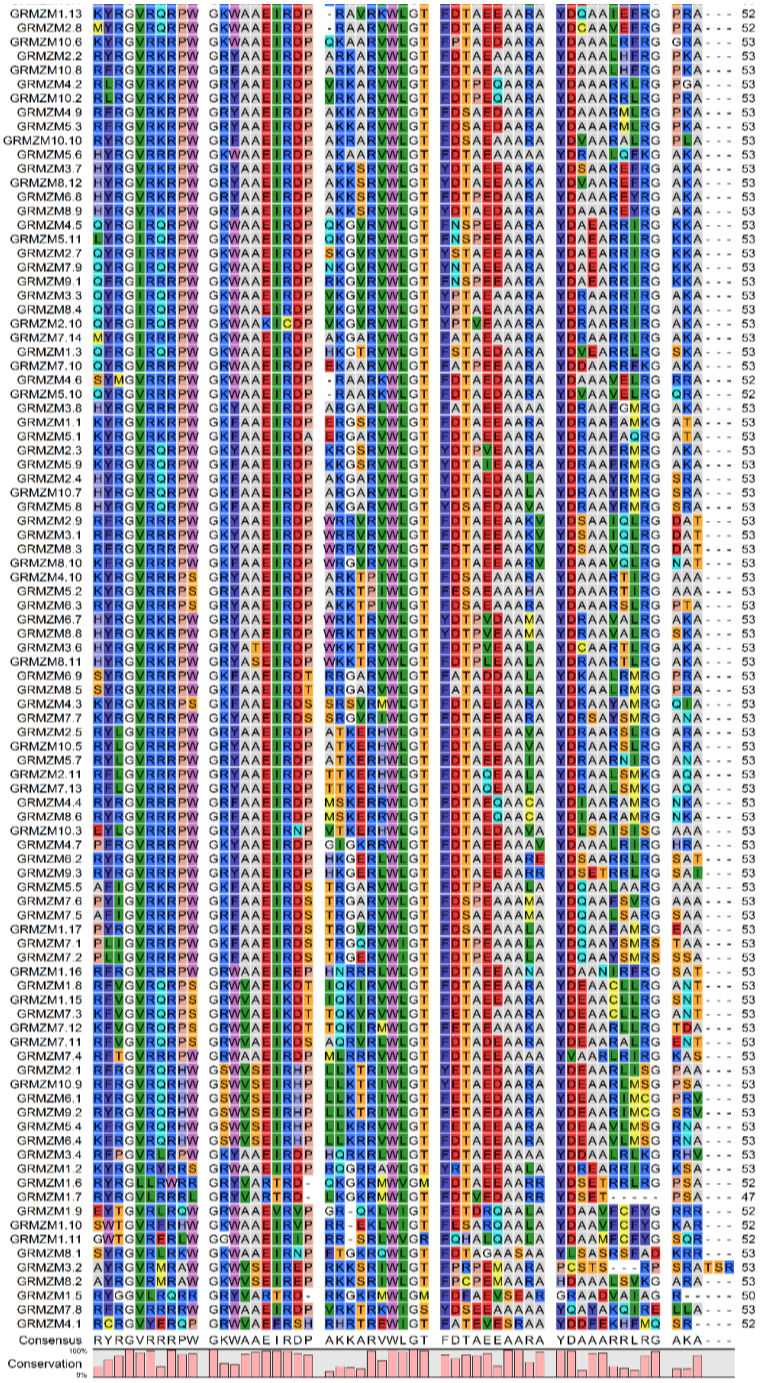


Figure 1. *Multiple sequence alignment of 105 ZmERF amino acid sequences, performed through CLC bio software package*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

Phylogenetic relationship between the ERF genes of Z. mays

Phylogenetic analysis revealed that 105 ERF genes are distributed into 38 sister groups and 29 genes are single gene. On the basis of phylogenetic analysis, ERF protein family is divided into 10 subgroups, namely I-X. Group I contain 4 ERF genes, II contain 7 genes, III 9 genes, IV has 12 genes, V has 9 genes, VI has 15 genes, VII have 9 genes, VIII have 15 genes, IX have 16 genes and X have 9 genes (*Fig. 2*). In order to better understand relative relationship of the same gene family among different species, an evolutionary analysis is performed between *Z. mays* (105 ERF genes) and Sorghum (53 ERF genes) (Yan et al., 2013), as both species belong to monocot and sorghum represented a model plant. Phylogenetic tree showed all *Z. mays* and sorghum ERF genes are distributed into 59 sister groups, containing 27 ZmERF-ZmERF, 28 ZmERF-SbERF and 4 SbERF-SbERF (*Fig. 3*).

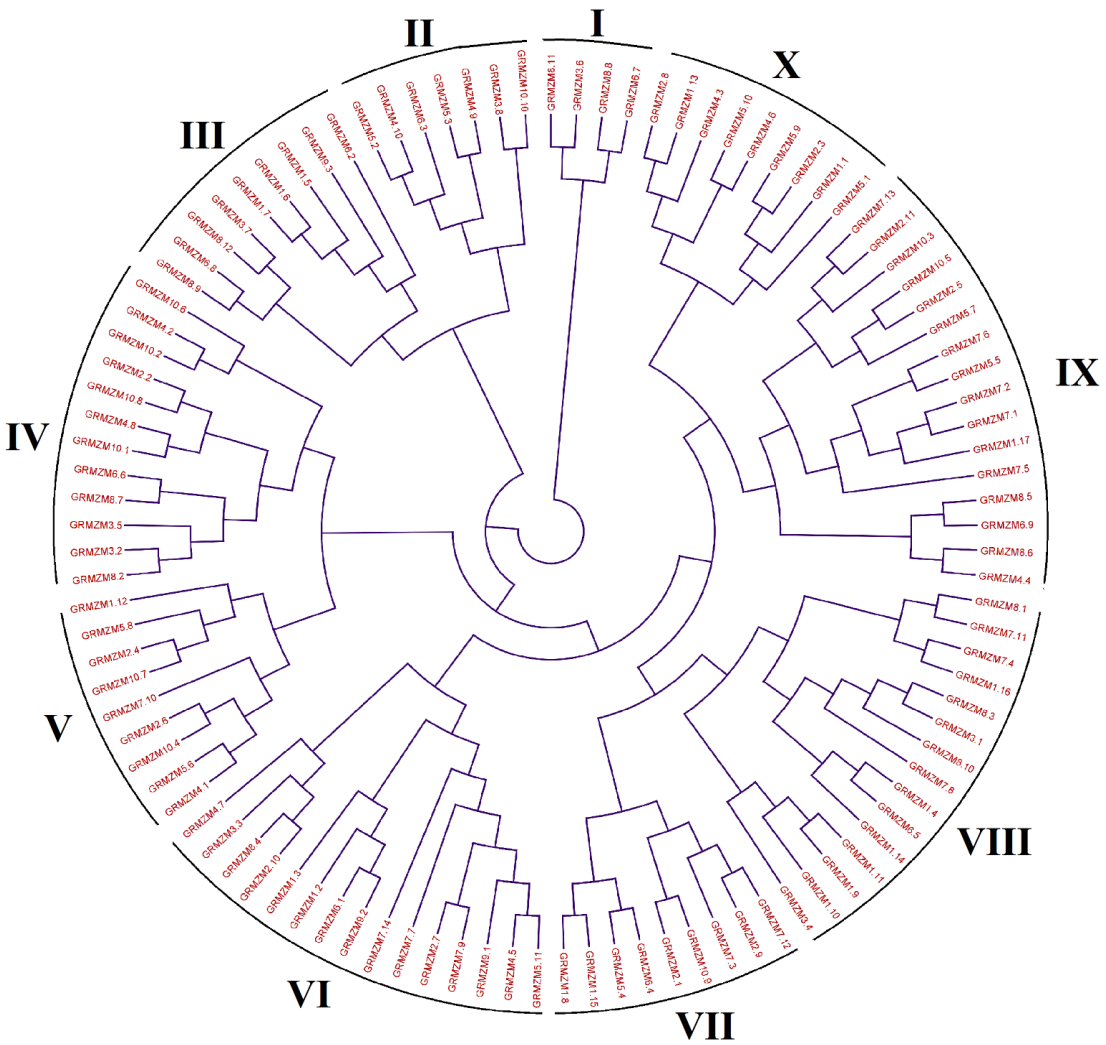


Figure 2. *Phylogenetic analysis of Z. mays ERF genes. An uprooted tree was generated using the CLC sequence viewer 7.0 program by the neighbor-joining method with bootstrap 1000 replicates*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

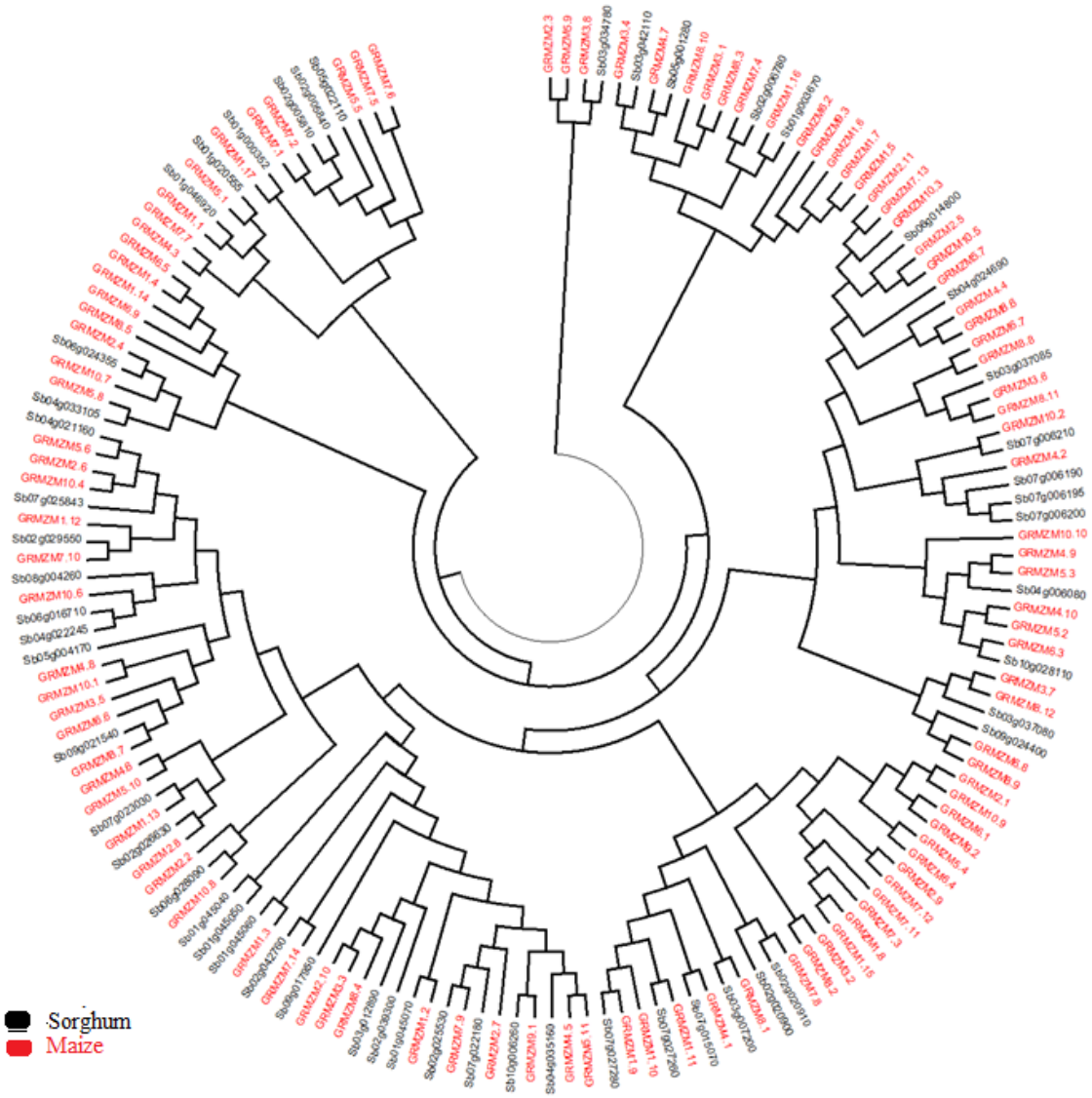


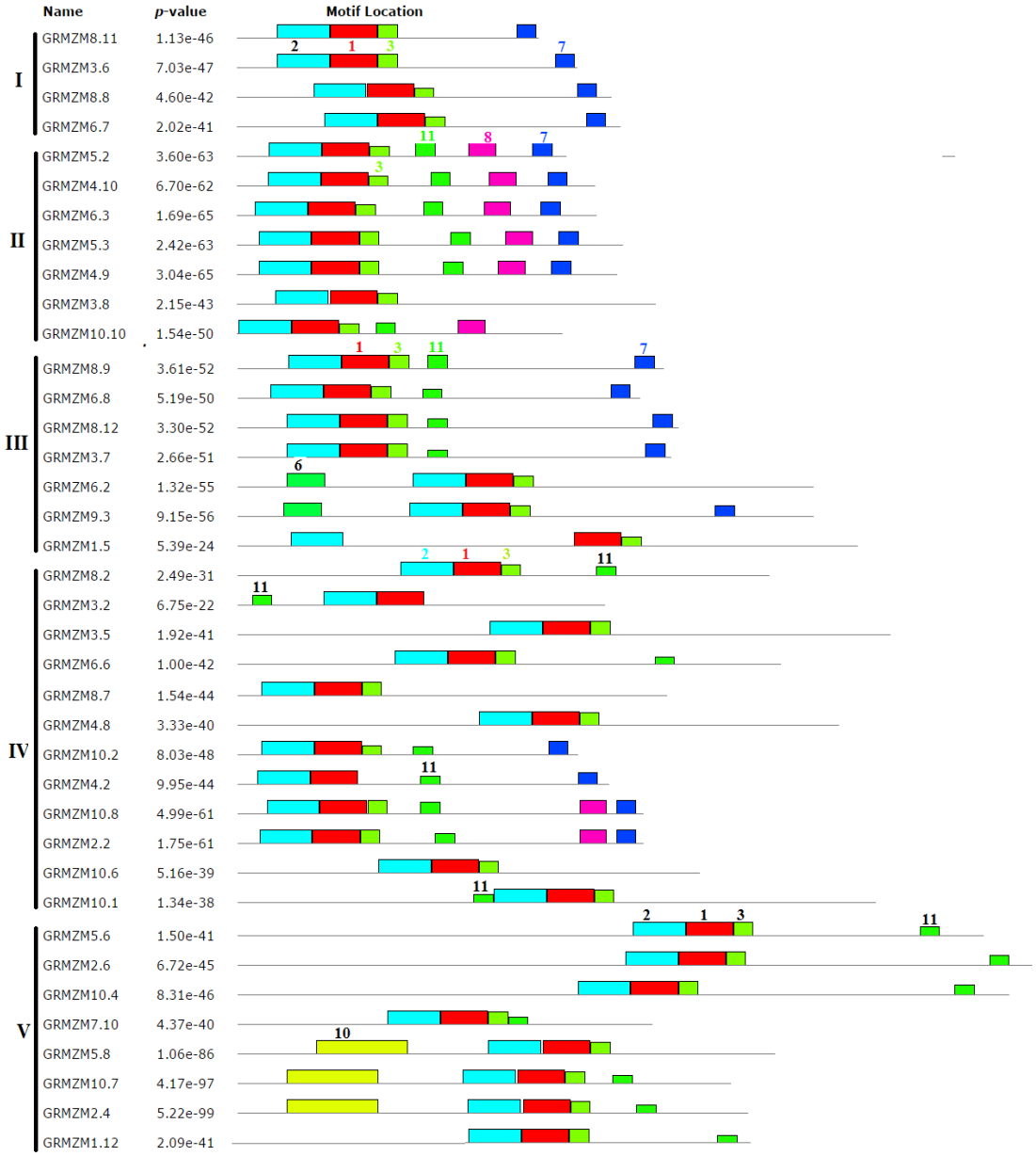
Figure 3. *An uprooted phylogenetic tree constructed by MEGA6 using neighber-joining method of Z. mays and sorghum ERF genes.*

Analysis of conserved motifs in ZmERF family

MEME motif discovery depicted that there were many other motifs then ERF motif. In *Figure 4*, motif 1, 2 and 3 were conserved in all 105 ERF protein sequence while other motifs were group specific. In 1st group Motif 7th was added at the N-terminal. Similarly in group II, motif 8th and 11th were observed in addition to the motifs in I group. 3rd group (III) contained motif 6th while there is deletion of motif 8th. In 4th group some gene had 7th motif in addition to 1st, 2nd, 3rd and 11th motifs. Group 5th had 10th motif with the motifs present in group I. IV group had motif 5th, motif 9th some proteins had motif 12th and 4th. In VII group, motif 14th is added with other motifs. In group VIII motif 6 is dominant with motifs 1, 2 and 3. In IX, motif 13th was present in addition to special motif, 4 and 11. In the last group(X) there were no any new motifs other than mentioned above (*Fig. 4*).

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

Some ZmERF protein sequences also showed LWSY(Motif 12) and EAR-like (Motif 7) motifs just like other plant species *viz. Arabidopsis*, rice and sorghum (Yan et al., 2013). In *Z. mays* LWSY motif is modified in LWSF pattern like in A-3 group of sorghum (Yan et al., 2013), and only three ERF genes have LWSF motif, GRMZM4.5, GRMZM5.11 and GRMZM9.1 belong to group VI. EAR-like motif with conserved residues DLNXP were present in 17 ERF genes covering group I, II and III. These motif along their gene name, start point and p-values are given in *Fig. 5*.



http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

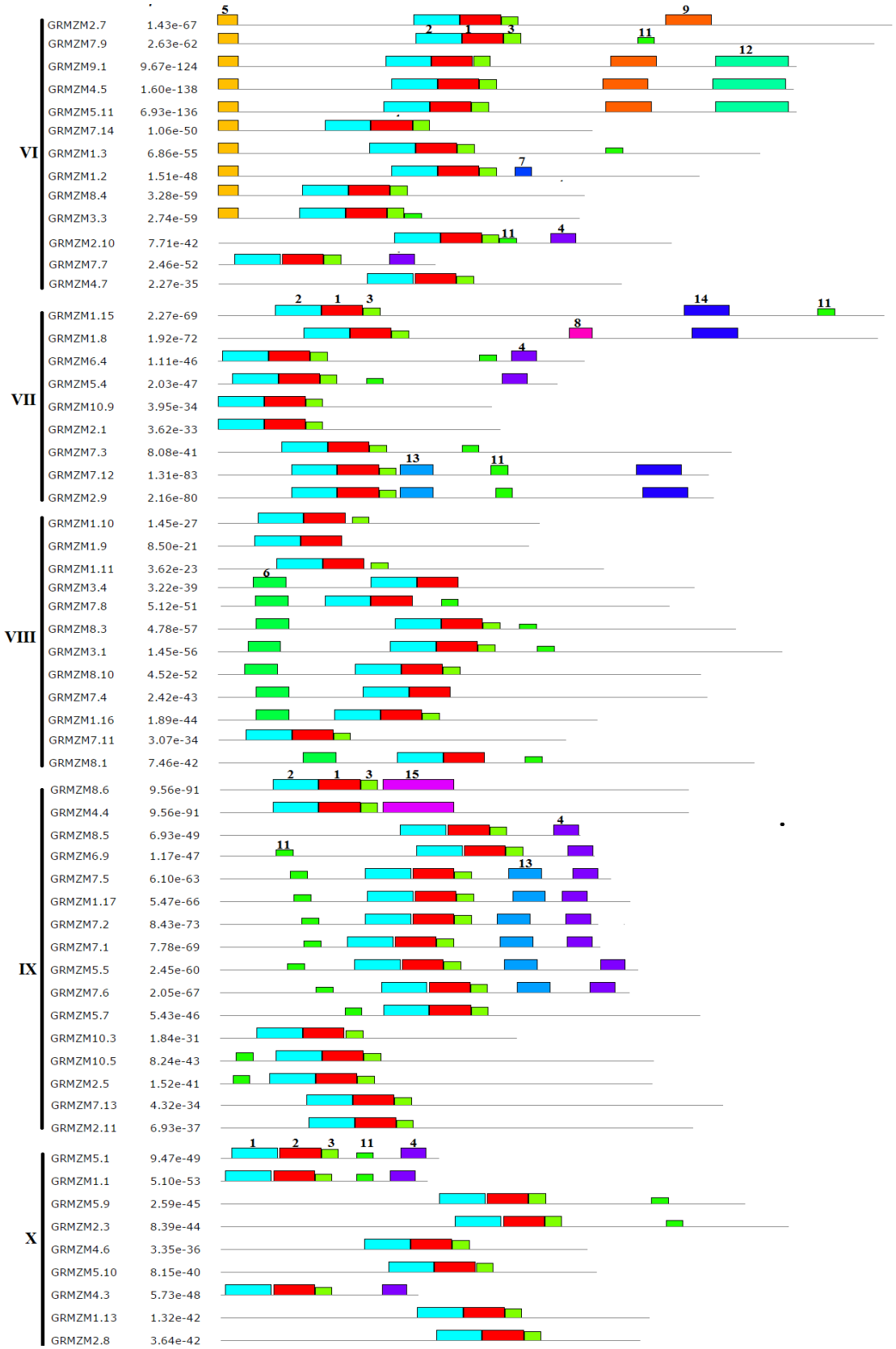


Figure 4. *Conserved Motif identified among group I-X of ERF family in Z. Mays, using MEME Motif discovery analysis. Gene name and combined E-velue are given at left side, each color represents specific motif and block length represent length of motif*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

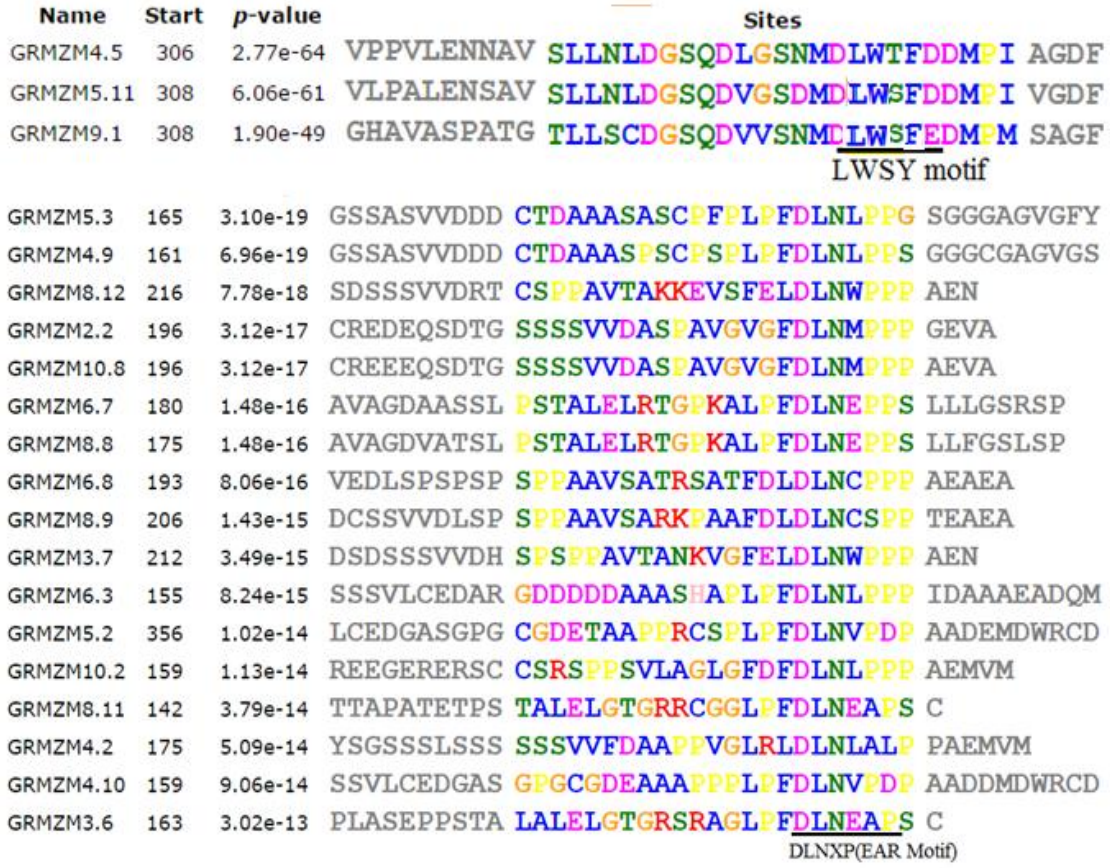


Figure 4. *Conserved LWSY motif and ERF-associated amphiphilic repression (EAR) motif sequences in the c-terminal region of ERF proteins. Consensus amino acid residues were identified using the MEME program and the conserved motifs are underlined respective motif 12 and Motif 7 respectively*

Mapping of ERF genes on Z. mays chromosomes

According to various sources data collection, chromosomal location of 105 genes was predicted graphically. Comparison showed chromosome 1 has highest number of gene (17 genes), followed by chromosome 7 and 8 (14 and 12 genes, respectively). Chromosome 2 and 5 had 11 genes, 4 and 10 carried 10 genes and 6, 3 and 9 harbor 9, 8, 3 genes, respectively. Result of chromosomal location showed most of the ZmERF genes were present in cluster form. Interestingly, almost all genes are either on top or bottom of the chromosome. Many genes in the same group are located on the one chromosome, for instance, GRMZM1.5 GRMZM1.6 GRMZM1.7 belongs to group III clustered in a ~34 kbp segment (chr1);.GRMZM7.1 and GRMZM7.2 belong to IX group, are clustered in ~16 kbp segment (chr7) (*Fig. 6*).

Gene structure analysis of ZmERF

In order to observe detailed evolutionary relationship among ZmERF genes, intron-exon structure was analyzed. Result of the current study showed that most of the genes have no intron and some of them have one or two introns. GRMZM1.7

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

gene has highest number of exons (5), followed by GRMZM1.6, GRMZM1.2 and GRMZM10.4 (4, 3, 3 exons, respectively) (*Fig. 7*).

Subcellular localization of ZmERF

In subcelular localizatoin analysis, most of ZmERF proteins were confined in nuclues. According to Plant-mPLoc prediction, 94 ZmERF genes were localized only in nuclues, two genes GRMZM8.1 and GRMZM8.2 were present in cytoplasm and 8 ZmERF genes shared multiple locations (cytoplasm, nucleus, cell membrane and chloroplast). As shown in *Table 3*, Wolfpsort predicted 22 ZmERF genes located in Nucleus, 42 proteins shared dual locations (nucleus and chloroplast) and some proteins were found at multi-located (Cytoplasm, Chloroplast, Nucleus, Mitochondria).

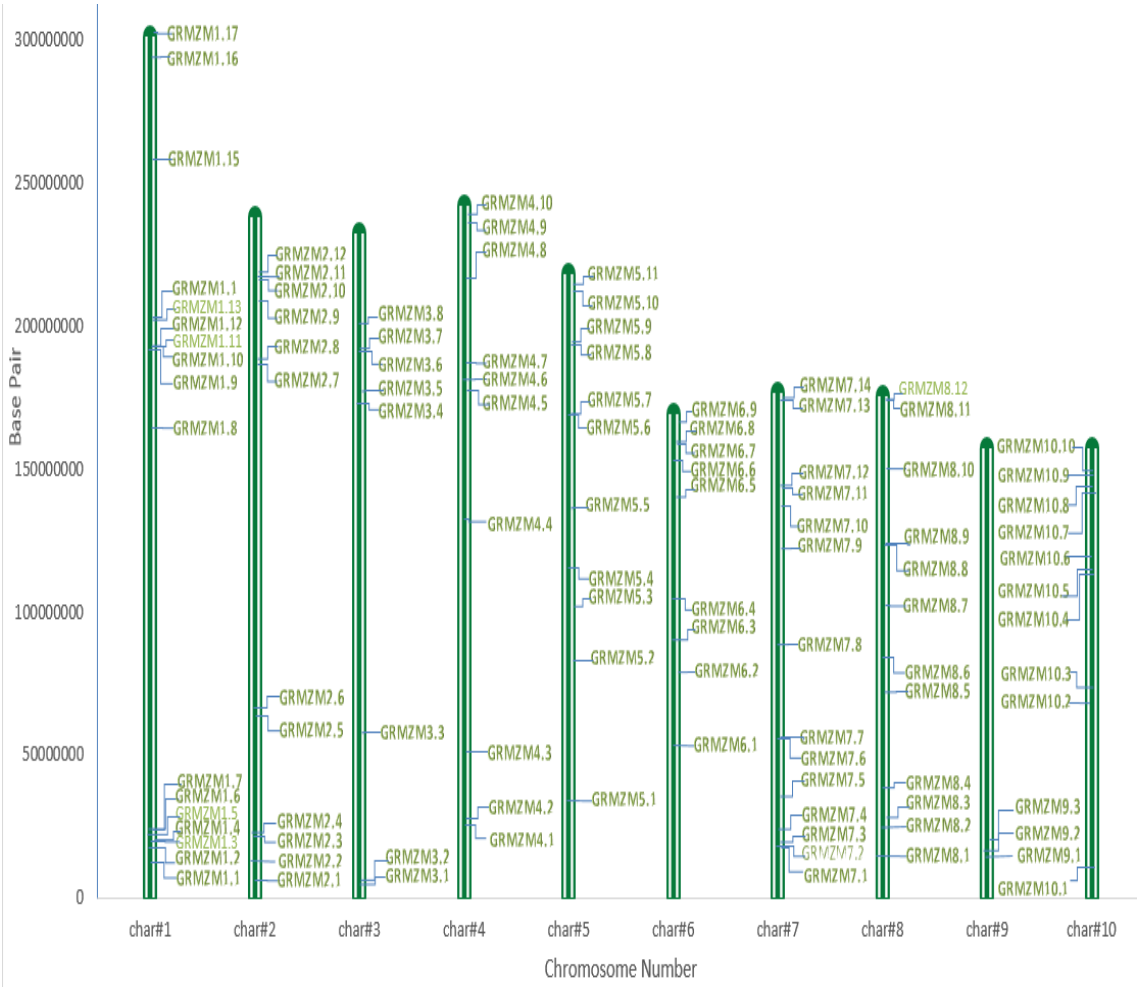


Figure 5. *Chromosomal mapping: The localization of 105 ZmERF genes on Z. mays chromosomes. The chromosomes number is shown at the bottom of each bar. ERF Genes are named according to their position and size on the chromosome and mentioned in the Table 2. The comparative position of ERF gene and size of chromosome are characterized using vertical scale.*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

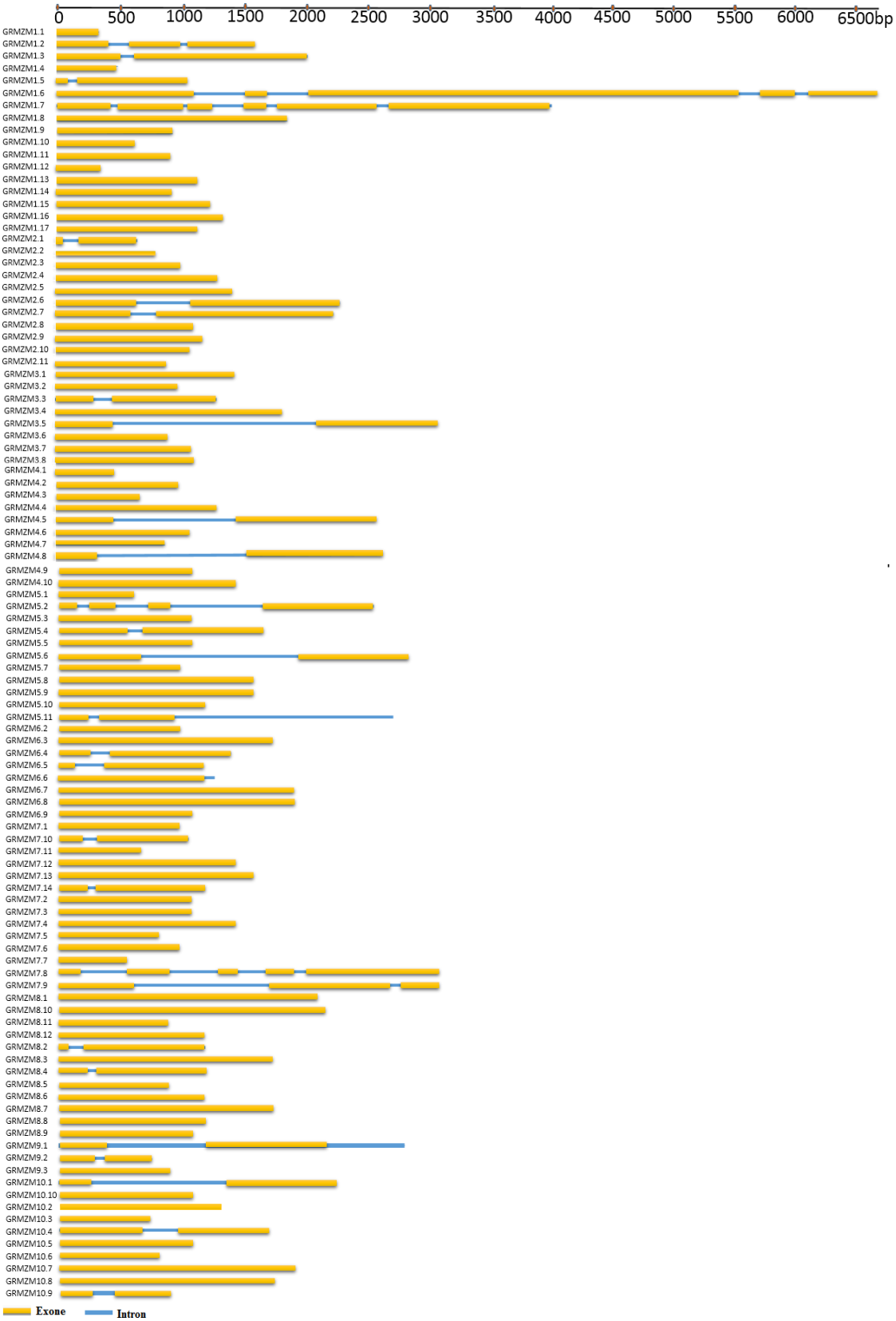


Figure 6. *Gene structure analysis: intron/exon structure ERF genes from Z. Mays. Scale is drawn according to largest gene. Brown color is showing exon portion of gene while blue line represent intron. The size of intron and exon are drawn according to the scale at the top*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

***Table 3**. Predicted subcellular localization of 105 ZmERF, with their annotation, molecular weight and isoelectric point*

Annotation	Mw (kDa)	pI	wolfpsort	Plant-mPLoc[1]
GRMZM1.1	13.89	8.46	nucl: 13	Nucleus
GRMZM1.2	32.69	5.15	chlo: 5, cyto: 4	Nucleus
GRMZM1.3	36.89	4.91	nucl: 13	Nucleus
GRMZM1.4	15.51	11.65	chlo: 7	Chloroplast. Nucleus
GRMZM1.5	36.43	7.18	nucl: 5, mito: 5, cyto: 4	Cytoplasm. Nucleus
GRMZM1.6	216.02	10.58	nucl: 14	Cytoplasm. Nucleus
GRMZM1.7	127.97	10.31	nucl: 12, cyto: 1	Cytoplasm. Nucleus
GRMZM1.8	43.94	4.37	chlo: 6, nucl: 5, cyto_nucl: 4,	Nucleus
GRMZM1.9	20.96	10.94	chlo: 11, nucl: 3	Cytoplasm. Nucleus
GRMZM1.10	21.97	4.83	nucl: 7, cyto: 4, chlo: 2	Cytoplasm. Nucleus
GRMZM1.11	25.49	7.93	chlo: 6, mito: 4,	Cytoplasm
GRMZM1.12	16.75	10.88	nucl: 12, plas: 1	Nucleus
GRMZM1.13	29.03	6.37	nucl: 9, chlo: 4	Nucleus
GRMZM1.14	18.9	10.82	cyto: 9, chlo: 3	Nucleus
GRMZM1.15	44.44	4.77	cyto: 7, chlo: 4,	Nucleus
GRMZM1.16	25.63	11.43	nucl: 9, cyto: 4	Nucleus
GRMZM1.17	37.6	9.95	chlo: 9, nucl: 5	Nucleus
GRMZM2.1	25.92	5.35	chlo: 5, mito: 4	Nucleus
GRMZM2.2	29.77	10.23	chlo: 13	Nucleus
GRMZM2.3	44.62	4.58	cyto: 8, nucl: 5	Nucleus
GRMZM2.4	36.72	5.18	nucl: 12, chlo: 2	Nucleus
GRMZM2.5	35.1	4.23	chlo: 8, nucl: 6	Nucleus
GRMZM2.6	54.52	9.94	chlo: 7, nucl: 4	Nucleus
GRMZM2.7	51.36	4.57	chlo: 9, nucl: 3	Nucleus
GRMZM2.8	34.76	6.28	chlo: 7, cyto: 5,	Nucleus
GRMZM2.9	40.35	8.12	nucl: 14	Cytoplasm. Nucleus
GRMZM2.10	37.49	9.05	chlo: 10, nucl: 3	Nucleus
GRMZM2.11	39	7.06	nucl: 9, chlo: 4	Nucleus
GRMZM3.1	45.03	4.49	nucl: 12, cyto: 1	Nucleus
GRMZM3.2	29.03	12.14	nucl: 13	Nucleus
GRMZM3.3	24.49	7.34	nucl: 11, chlo: 2	Nucleus
GRMZM3.4	31.73	7.58	chlo: 9, nucl: 5	Cytoplasm. Nucleus
GRMZM3.5	36.75	6.24	nucl: 14	Nucleus
GRMZM3.6	19.65	10.73	chlo: 6, cyto: 3	Nucleus
GRMZM3.7	24.43	9.02	chlo: 13	Nucleus
GRMZM3.8	23.29	9.12	nucl: 10, chlo: 4	Nucleus
GRMZM4.1	18.51	7.64	nucl: 6, mito: 5, cyto: 2	Nucleus

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

GRMZM4.2	20.76	7.35	chlo: 8, nucl: 4, mito: 2	Nucleus
GRMZM4.3	13.24	9.09	nucl: 8, mito: 5	Nucleus
GRMZM4.4	31.34	8.30	nucl: 12, cyto: 2	Nucleus
GRMZM4.5	39.47	4.44	nucl: 13	Nucleus
GRMZM4.6	23.97	8.50	nucl: 10, chlo: 3	Nucleus
GRMZM4.7	27.05	7.87	chlo: 7, cyto: 3	Nucleus
GRMZM4.8	34.31	730	nucl: 13	Nucleus
GRMZM4.9	21.69	8.14	nucl: 10, mito: 3	Nucleus
GRMZM4.10	19.75	5.29	nucl: 7, chlo: 4	Nucleus
GRMZM5.1	14.1	7.53	nucl: 14	Nucleus
GRMZM5.2	42.58	9.48	chlo: 7, nucl: 5	Nucleus
GRMZM5.3	22.16	7.88	nucl: 10, chlo: 3	Nucleus
GRMZM5.4	23.36	10.03	nucl: 13	Nucleus
GRMZM5.5	30.61	6.40	nucl: 10, cyto: 2	Nucleus
GRMZM5.6	43	8.86	nucl: 12, chlo: 2	Nucleus
GRMZM5.7	32.81	4.40	nucl: 13	Nucleus
GRMZM5.8	30.72	5.81	nucl: 12	Nucleus
GRMZM5.9	34.51	6.10	nucl: 13	Nucleus
GRMZM5.10	25.02	8.94	nucl: 7, chlo: 5	Nucleus
GRMZM5.11	39.63	4.69	nucl: 14	Nucleus
GRMZM6.1	18.81	6.96	mito: 10, chlo: 3	Nucleus
GRMZM6.2	33.16	4.50	nucl: 14	Nucleus
GRMZM6.3	20.61	6.35	nucl: 6, mito: 5	Nucleus
GRMZM6.4	23.83	10.47	chlo: 7, nucl: 4, mito: 3	Nucleus
GRMZM6.5	24.08	13.03	chlo: 6, nucl: 6, mito: 2	Nucleus
GRMZM6.6	31.26	8.57	nucl: 9.5, cyto_nucl: 6,	Nucleus
GRMZM6.7	21.73	10.92	cyto: 6, nucl: 4,	Nucleus
GRMZM6.8	23	8.20	chlo: 10, mito: 3	Nucleus
GRMZM6.9	31.03	6.94	nucl: 14	Nucleus
GRMZM7.1	35.08	6.34	nucl: 14	Nucleus
GRMZM7.2	38.19	5.79	nucl: 14	Nucleus
GRMZM7.3	34.37	9.64	nucl: 12, mito: 2	Nucleus
GRMZM7.4	33.2	10.06	nucl: 13	Nucleus
GRMZM7.5	27.06	10.25	nucl: 7, chlo: 5	Nucleus
GRMZM7.6	27.61	6.51	nucl: 13	Nucleus
GRMZM7.7	15.19	8.45	nucl: 13	Nucleus
GRMZM7.8	31.47	4.42	nucl: 7, chlo: 4	Nucleus
GRMZM7.9	44.16	5.17	chlo: 12, nucl: 1	Nucleus
GRMZM7.10	23.92	9.79	nucl: 11, mito: 2	Nucleus
GRMZM7.11	22.97	10.42	chlo: 11, mito: 2	Nucleus

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

GRMZM7.12	33.03	7.18	nucl: 14	Nucleus
GRMZM7.13	33.34	7.11	nucl: 13	Nucleus
GRMZM7.14	25.74	6.29	nucl: 12, cyto: 2	Nucleus
GRMZM8.1	35.99	4.75	chlo: 10, nucl: 2	Cytoplasm
GRMZM8.2	29.78	7.68	chlo: 6, nucl: 3, mito: 3	Cytoplasm
GRMZM8.3	35.77	4.43	nucl: 7, cyto: 3, chlo: 2	Nucleus
GRMZM8.4	24.77	8.87	nucl: 12, chlo: 1	Nucleus
GRMZM8.5	30.14	7.59	nucl: 13	Nucleus
GRMZM8.6	31.34	8.30	nucl: 12, cyto: 2	Nucleus
GRMZM8.7	24.19	6.97	nucl: 10.5, nucl_plas: 6	Nucleus
GRMZM8.8	21.62	10.69	cyto: 6, nucl: 4	Nucleus
GRMZM8.9	23.88	8.46	chlo: 11, nucl: 2	Nucleus
GRMZM8.10	31.77	4.41	nucl: 12, chlo: 2	Nucleus
GRMZM8.11	17.84	10.29	nucl: 7, chlo: 6	Nucleus
GRMZM8.12	24.94	9.36	chlo: 14	Nucleus
GRMZM9.1	39.31	4.97	chlo: 9, nucl: 4	Nucleus
GRMZM9.2	18.03	6.26	chlo: 6, mito: 4, nucl: 3.5	Nucleus
GRMZM9.3	32.99	4.75	chlo: 13	Nucleus
GRMZM10.1	35.68	9.42	nucl: 9.5, plas: 5.5	Nucleus
GRMZM10.2	19.91	7.92	nucl: 13	Nucleus
GRMZM10.3	19.84	7.20	nucl: 10, mito: 4	Nucleus
GRMZM10.4	45.36	10.10	nucl: 14	Nucleus
GRMZM10.5	28.37	4.35	chlo: 9, nucl: 5	Nucleus
GRMZM10.6	26.06	9.88	nucl: 10, pero: 4	Nucleus
GRMZM10.7	29.01	5.73	nucl: 11.5, plas: 6.5	Nucleus
GRMZM10.8	22.83	10.52	chlo: 13	Nucleus
GRMZM10.9	18.38	6.67	nucl: 6.5, mito: 5, cyto_nucl: 4, chlo: 2	Nucleus
GRMZM10.10	18.94	10.30	nucl: 7, chlo: 5, mito: 2	Nucleus

nucl=nucleus, chlo = chloroplast, mito = mitochondria, cyto = cytoplasm

Gene ontology annotation

Gene ontology analysis resulted in the association of ZmERF proteins in diverse biological, cellular, and molecular activities (*Fig. 8*). The analysis of physiological processes connected to these ZmERF proteins revealed that most of the proteins were involved in stresses responses and regulations. Stress responses include water stress, salt stress, abiotic, hormone, lipid , cold, alcohol, while, in regulation processes, all ZmERF proteins were found to involve in primary metabolic process, nitrogen compound metabolism, macromolecule metabolism, cellular metabolism, biosynthetic process. In molecular process, all ZmERF proteins showed that all proteins possess nucleic acid binding activity including, catalytic activity (found in

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

one gene only) and cellular transportation (found in 2 genes). Moreover, cellular compartment by Blast2GO also predicted the localization of ZmERF proteins in nucleus (99 ZmERF genes), followed by plastids (10 ZmERF genes), mitochondria (5 ZmERF genes), and other intracellular membrane bounded organelles. In addition species distribution determination through Blast2GO tool resulted that *Oryza sativa* has highest blast hits with ZmERF proteins (~272 Blast hits value), followed by *Setaria italica* (~205 hits), *Sorghum bicolor* (~128 hits) and lowest blast hits is shown by *Solanum lycopericum* (~19 blast hits) (*Fig. 9*).

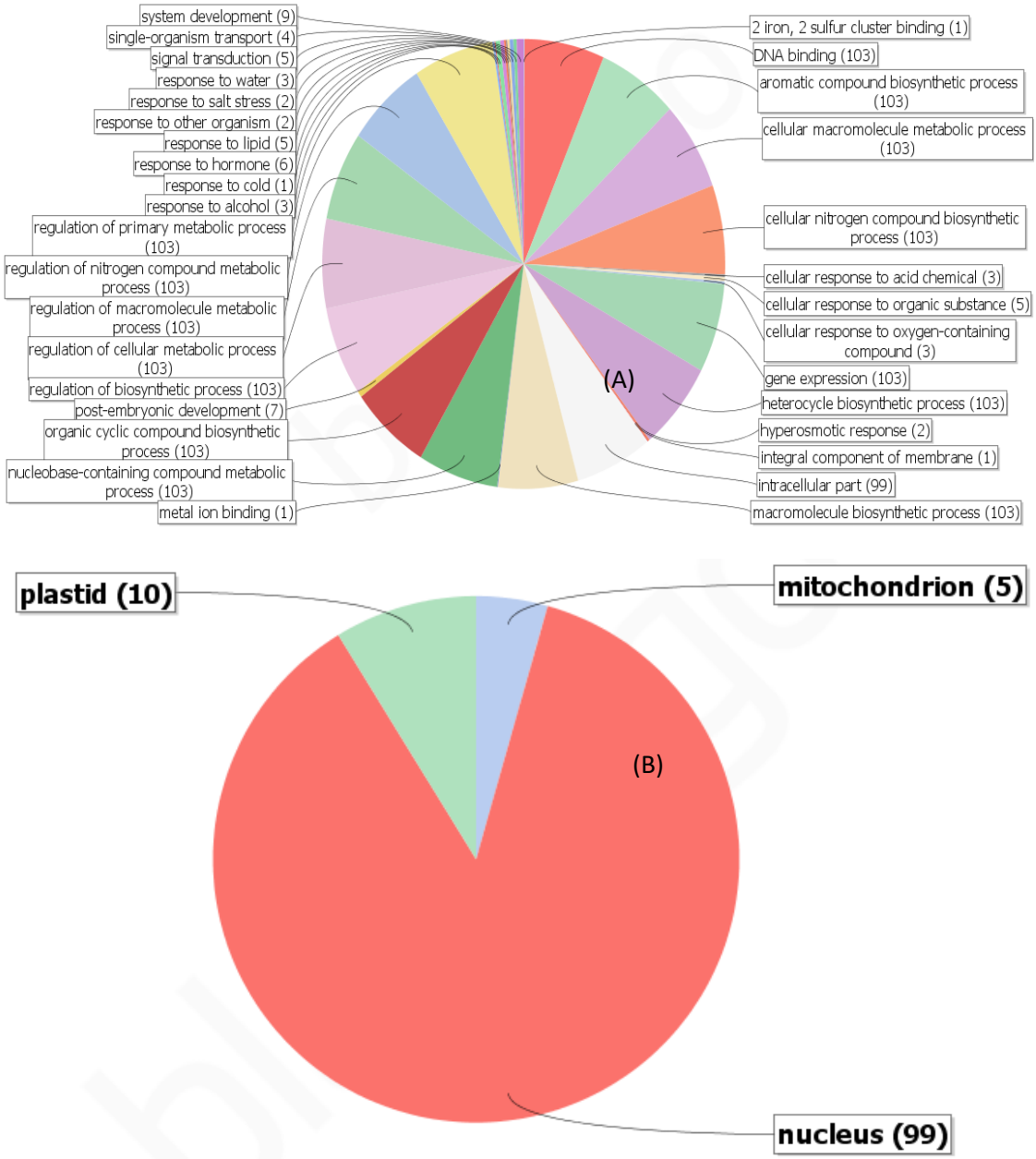


Figure 7. *Gene Ontology of ZmERF proteins .The result of Blast2GO showing (A) Biological Process and (B) Cellular Component , of 105 ZmERF proteins*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

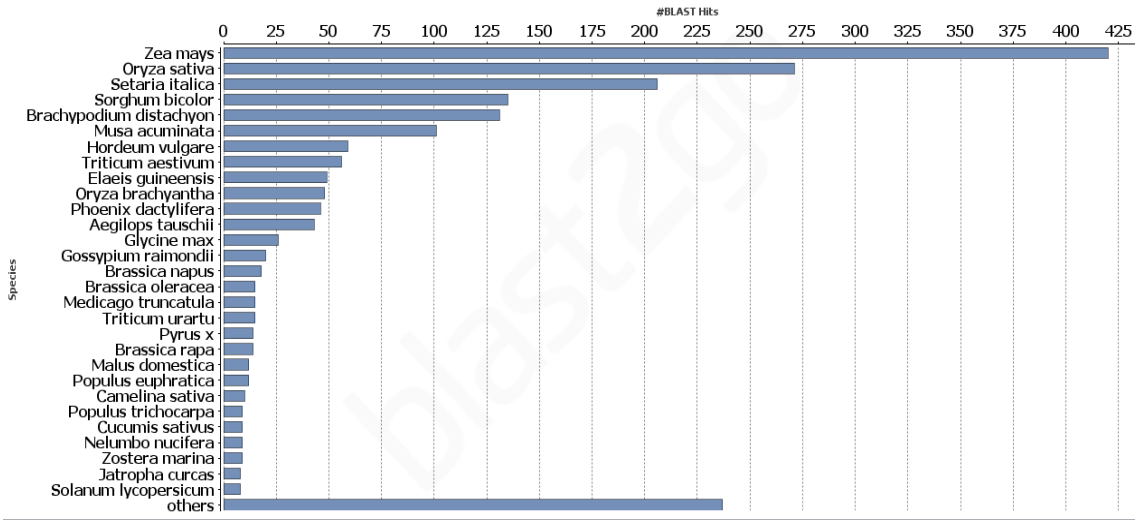


Figure 9. *Species distribution of top BLAST hits obtained using Z. Mays amino acid residues BLAST hits against NCBI-Nt and other default databases in Blast2GO tool , are in x-axis and species distribution are on y-axis*

Discussion

Transcription factors play an important role in modulating the adaptation response of plants to various internal or external signals (Sharma et al., 2010). They possibly change downstream gene expression in stress signal transduction pathways via activation and repression of genes after experience to stress. Plant genomes comprise a large number of transcription factors like AP2/ERF, WORKY, MYB, SPL ARF, FAR1 (Zheng et al., 2007; Park et al., 2010; Chen et al., 2012; Ambawat et al., 2013; Padmanabhan et al., 2013; Mathelier et al., 2014). It has been expected that the *Arabidopsis* genome codes for at least 1533 transcription factors, comprising of over 5.9% of its total predicted genes (Rao et al., 2014). According to (Du et al., 2014), 105 AP2/ERF genes were identified in ERF subfamily having one AP2-like domain.

Throughout the last decade, a huge amount of research has been directed that showed overexpression of ERF family genes increases the resistance of plants to environmental stresses (Shinozaki and Shinozaki, 2012). For example, overexpression of ERF genes in chillies (Shin et al., 2002), *Arabidopsis* (Berrocal-Lobo et al., 2002), and tomato (Gu et al., 2002). Therefor it is one of the most important challenges to identify the genes which have resistant against abiotic factor. The phylogeny, conserved motifs, gene structure and gene ontology annotation were extensively analyzed in several plant species including genome-wide analysis in model plant, *Arabidopsis*, rice (Nakano et al., 2006) and sorghum (Yan et al., 2013), while few ERF gene is *Z. mays* were also studied in waterlogging stress (Liu et al., 2008). With the completion of *Z. mays* genome sequencing projects, phylogenetic analysis of ZmERF genes would be helpful to investigate general function of ERF genes and the evolutionary process of the AP2/ERF domain in plants.

We found that *Z. mays* has 105 ERF genes which are greater than *Arabidopsis*, sorghum, peach and moso bamboo, while less than Chines cabbage and Carrot (*Table 1*). Within conserved domain, G4, G11, W27 and G29 are conserved in all 105 ERF genes while R6, R8, A15, E16, I17, R18, R25, L28, T30, A37, A38, Y41, D41 and G50

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

are conserved in 97% of 105 ERF genes (*Figure 1*). A similar pattern of conserved residues was reported in sorghum ERF genes (Yan et al., 2013) and in cucumbers (Hu and Liu, 2011). These conserved residues have significant effect in the function of AP2-domain and also help for point mutation on individual protein. Comparative analysis of amino acid residues among different crops have shown that the most conserved motifs in the AP2/ERF superfamily were present in other species, including *Arabidopsis*, rice, maize and other plants like LWSF (Motif 12) and EAR (Motif 7) (Yan et al., 2013) are also observed in ZmERF proteins in which LWSF function was in gene repression transcriptional regulatory cascades (Ohta et al., 2001). Apart from commonly present motifs in *Arabidopsis* and rice, other motifs in *Z. mays* also play important role in biological processes. In contrast to the groupings (B1-B6) described by (Nakano et al., 2006), we have classified ERF genes in to ten groups (I-X) which will provide more detailed information, where IX had largest number followed by VI and VIII. A relatively great number of genes in these groups might be the significance of evolutionary adaptations to numerous environmental changes. The 105 ZmERF have found 39 sister pairs which showed strong nitration and duplication among ZmERF proteins. The comparative analysis of ZmERF with sorghum ERF showed 28 ZmERF-SbERF sister pair with close homology in ERF family among monocots species.

In chromosomal maps, distribution of ZmERF genes in all the 10 chromosomes was similar to the findings of AP2/ERF in Sorghum (Yan et al., 2013) and Chinese cabbage (Song et al., 2013). Interestingly, ERF gene comprising of same group and present on same chromosome forming cluster, was also observed in *Arabidopsis* (Sakuma et al., 2002) sorghum (Yan et al., 2013), brassica (Song et al., 2013) and cucumbers (Hu and Liu, 2011). The clusters of genes with same product function are useful for drawing recent evolutionary history (Anollés, 2001).

It has been found that intron/exon position configuration provides hint in evolutionary relationship (Hu and Liu, 2011). A very limited number of introns were found in ZmERF genes, while most of the genes were dominated by exons. In previous studies, it has been reported that most of the ERF genes in *Arabidopsis* and peach (Sakuma et al., 2002; Zhang et al., 2012) have no introns. Current schematics representation showed that the absence of intron in ERF genes is also a feature of *Z. mays* (*Table 2*; *Fig. 6*).

Out of 105, 103 ZmERF genes were annotated by Blast2GO. Annotation resulted in the putative participation of ZmERF in diverting the biological, cellular and chemical processes including abiotic stress responses like drought, cold, temperature, salinity etc. Our finding is similar to previous study of AP2/ERF genes in other species, like expression of *PeDREB2* in *Populus euphratica* carried enhancement of drought and cold (Chen et al., 2009; Kumar and Venkateswarlu, 2011) and expression of Sl-ERF.B.3 gene in tomato was only observed under stress condition (Klay et al., 2014). The molecular process of ZmERF showed that all 105 genes have sequence-specific DNA-binding activity which is already reported as a feature of AP2/ERF gene (Yamasaki et al., 2013) and also studied in cucumbers (Hu and Liu, 2011). Cellular localization showed 99 genes are located in nucleus, 10 in plastid and 5 in mitochondria. This means that some ZmERF have multiple locations in the cell. This characteristic of AP2/ERF genes was reported in foxtail millet (Lata et al., 2014).

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

Conclusion

To our knowledge, this is the first study in *Z. mays* regarding genome-wide analysis of ERF genes. The study elucidated the ERF gene family role in regulations and defense responses and stress signaling pathways. Characterization and analysis of these functional TF genes may contribute to improve molecular basis of *Z. mays* for better gene pool development and stress adaptation. Our study may also attribute functional gene resources for genetic future engineering approaches for making the *Z. mays* plants tolerant again abiotic stresses. However, further research is needed to explore the biological roles of these TF genes.

Acknowledgement. We are thankful to Mr. Sakhawat Riaz, Department of Food, Nutrition and Technology, Government College University Faisalabad for helping in data arrangement and analysis.

REFERENCES

[1] Alonso, J.M. (2003): Genome-Wide Insertional Mutagenesis of *Arabidopsis thaliana.* - Science 301: 653–57.
[2] Ambawat, S., Sharma, P., Neelam, R., Yadav, S., Yadav, R.C. (2013): MYB Transcription factor genes as regulators for plant responses: An overview. - Physiology and Molecular Biology of Plants 19: 307–21.
[3] Andorf, M. et al. (2015): Maize GDB update: New tools, data and interface for the maize model organism database. - Nucleic Acids Research 5: 1007-1009.
[4] Bailey, T.L., Elkan, C. (1994): Fitting a mixture model by expectation maximization to discover motifs in biopolymers. proceedings / international conference on intelligent systems for molecular biology. - ISMB International Conference on Intelligent Systems for Molecular Biology 2: 28–36.
[5] Bailey, T.L., Williams, N., Misleh, C., Li., W.W. (2006): MEME: Discovering and analyzing DNA and protein sequence motifs. - Nucleic Acids Research 34: 69–73.
[6] Boutilier, K., Offringa, R., Sharma, V. K., Kieft, H., Ouellet, T., Zhang, L. (2002): Ectopic expression of baby boom triggers a conversion from vegetative to embryonic growth. - The Plant Cell 14: 1737–49.
[7] Boyer, J.S. (1982): Plant productivity and environment. - Science 218: 443–48.
[8] Anollés, G. (2001): Novel strategies to study the role of mutation and nucleic acid structure in evolution. - Plant Cell, Tissue and Organ Culture 67: 115–32.
[9] Chen, G., Jinhuan, L., Xinli, X., Weilun, Y. (2009): Expression profiling and functional characterization of a DREB2-type gene from populus euphratica. - Biochemical and Biophysical Research Communications 378: 483–87.
[10] Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., Yu, D. (2012): The role of WRKY transcription factors in plant abiotic stresses. - Gene Regulatory Mechanisms 1819: 120–28.
[11] Chou, C. K., Shen, H.B. (2010): Plant-mPLoc: A top-down strategy to augment the power for predicting plant protein subcellular localization. - Plos One 5: e11335.
[12] Conesa, A., Stefan, G. (2008): Blast2GO: A comprehensive suite for functional analysis in plant genomics. - International Journal of Plant Genomics: 1–12.
[13] Du, H., Huang, M., Zhang, Z., Cheng, S. (2014): Genome-wide analysis of the AP2/ERF gene family in maize waterlogging stress response. - Euphytica 198: 115–26.
[14] Dubouzet, J.G., Sakuma, Y., Ito, Y., Kasuga, M., Dubouzet, E.G., Miura, M. (2003): OsDREB genes in rice, *Oryza Sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. - Plant Journal 33: 751–63.
[15] Elliott, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W.Q. (1996): AINTEGUMENTA, an APETALA2-like gene of *Arabidopsis* with pleiotropic roles in

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

ovule development and floral organ growth. - The Plant cell 8: 155–68.
[16] Gong, W., He, K., Covington, M., Kumar, S.P., Snyder, M., Harmer, N. (2008): The development of protein microarrays and their applications in dna-protein and protein-protein interaction analyses of *Arabidopsis* transcription factors. - Molecular plant-microbe Intercation 1: 27–41.
[17] Gotz, S., Arnold, R., Sebastián-León, P., Martín-Rodríguez, S., Tischler, P. (2011): B2G-FAR, a species-centered GO annotation repository. - Bioinformatics 27: 919–24.
[18] Gu, Y.Q., Wildermuth, M.C., Chakravarthy, S., Loh, Y.T., Yang, C., He, X. (2002): Tomato transcription factors pti4, pti5, and pti6 activate defense responses when expressed in *Arabidopsis.* - The Plant cell 14: 817–31.
[19] Guo, Z.J., Chen, X.J., Wu, X.L., Ling, J.Q., Xu, P. (2004): Overexpression of the AP2/EREBP transcription factor OPBP1 enhances disease resistance and salt tolerance in tobacco. - Plant Molecular Biology 55: 607–18.
[20] Hao, D. , Yamasaki, K., Sarai, A., Ohme-Takagi, M. (2002): Determinants in the sequence specific binding of two plant transcription factors, CBF1 and NtERF2, to the DRE and GCC Motifs. - Biochemistry 41: 4202–4208.
[21] Hattori, Y., Nagai, K., Furukawa, S., Song, X.J., Kawano, R., Sakakibara, P. (2009): The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water - Nature 460: 1026–1030.
[22] Hinz, M., Wilson, I.W., Yang, J., Buerstenbinder, K., Llewellyn, D., Dennis, S. (2010): *Arabidopsis* RAP2.2: An ethylene response transcription factor that is important for hypoxia survival. - Plant Physiology 153: 757–72.
[23] Horton, H.P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. (2007): WoLF PSORT: protein localization predictor. - Nucleic Acids Research 35: 585–87.
[24] Hu, L., Liu, S. (2011): Genome-wide identification and phylogenetic analysis of the ERF gene family in cucumbers. - Genetics and Molecular Biology 34: 624–33.
[25] Hu, X.Y., Wang, Y.H., Liu, X.F., Li, Y.J. (2004): *Arabidopsis* RAV1 is down-regulated by brassinosteroid and may act as a negative regulator during plant development. - Cell Research 14: 8–15.
[26] Jiang, C., Iu, B., Singh, J. (1996): Requirement of a CCGAC cis-acting element for cold induction of the BN115 gene from winter *brassica napus.* - Plant molecular biology 30: 679–84.
[27] Jin, J., Zhang, H., Kong, L., Gao, G., Luo, J. (2014): PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. - Nucleic Acids Research 42: 1182–87.
[28] Jofuku, K.D., Boer, B.G., Montagu, M.V., Okamuro, J.K. (1994): Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. - The Plant cell 6: 1211–25.
[29] Kagaya, Y. Yasuaki, K., Hattori, K. (1999): RAV1, a novel dna-binding protein, binds to bipartite recognition sequence through two distinct dna-binding domains uniquely found in higher plants. - Nucleic Acids Research 27: 470–78.
[30] Kim, I.J.S., Mizoi, J., Yoshida, T., Fujita, Y., Nakajima, J., Ohori, T. (2011): An ABRE promoter sequence is involved in osmotic stress-responsive expression of the dreb2a gene, which encodes a transcription factor regulating drought-inducible genes in *Arabidopsis.* - Plant and Cell Physiology 52: 2136–46.
[31] Klay, I., Pirrello, J., Riahi, L., Bernadac, A., Cherif, A., Bouzayen, M. (2014): Ethylene response factor Sl-ERF.B.3 is responsive to abiotic stresses and mediates salt and cold stress response regulation in tomato. - The Scientific World Journal 167681.
[32] Knudsen, B., Knudsen, T., Flensborg, M., Sandmann, H., Heltzen, M., Andersen, A., Mønsted, S. (2011): CLC Sequence Viewer. A/S Cb 6.
[33] Kumar, A., Venkateswarlu, B. (2011): Abiotic stress response in plants-physiological , biochemical and genetic perspectives. - InTech, Rijeka Croatia: 1-358
[34] Lakhwani, D., Pandey, A., Dhar, Y.V., Bag, S.K., Trivedi, P.K., Asif, M. (2016):

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

genome-wide analysis of the AP2/ERF family in musa species reveals divergence and neofunctionalisation during evolution. - Scientific Reports 6: 18878.

[35] Lata, C., Mishra, A.K., Muthamilarasan, M., Bonthala, V.S., Khan, Y., Prasad, M. (2014): Genome-wide investigation and expression profiling of ap2/erf transcription factor superfamily in foxtail millet (*Setaria italica* L.). - PLoS ONE 9: e113092.

[36] Francesco, L., Francesco, M.G., Sara, Z., Fabio, O., Mario, P., Pierdomenico, P. (2010): Genomic and transcriptomic analysis of the AP2/ERF superfamily in vitis vinifera. - BMC genomics 11: 719.

[37] Xihui, L., Junji, F., Dan, G., Wenxin, L., Tingsong, L., Yunling, P., Jianhua, W., Gouying, W. (2008): Genome-wide analysis of gene expression profiles during the kernel development of maize (*Zea mays* L.). - Genomics 91: 378–87.

[38] Magnani, E., Sjölander, K., Hake, S. (2004): From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. -The Plant cell 16: 2265–77.

[39] Mathelier, A., Zhao, X., Zhang, A.W., Parcy, F., Worsley-Hunt, R., Arenillas, D. J., et al. (2014): JASPAR: An extensively expanded and updated open-access database of transcription factor binding profiles. - Nucleic Acids Research 42: 142–47.

[40] Mizoi, J., Shinozaki K., Yamaguchi-Shinozaki, K. (2012): AP2/ERF family transcription factors in plant abiotic stress responses. - Biochimica et Biophysica Acta-Gene Regulatory Mechanisms 1819: 86–96.

[41] Nakano, T., Suzuki K., Fujimura T., Shinshi H. (2006): Genome-wide analysis of the erf gene family. - Plant Physiology 140: 411–32.

[42] Nole-Wilson, S., Krizek, B.A. (2000): DNA binding properties of the *Arabidopsis* floral development protein aintegumenta. - Nucleic Acids Research 28: 4076–82.

[43] Ohme-Takagi, M., Shinshi H. (1995): Ethylene-Inducible DNA binding proteins that interact with an ethylene-responsive element. - The Plant cell 7: 173–82.

[44] Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H., Ohme-Takagi, M. (2001): Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. - Plant Cell 13: 1959–68.

[45] Pabo, C.O., Sauer R.T. (1992): Transcription factors: structural families and principles of dna recognition. - Annual Review of Biochemistry 61: 1053–95.

[46] Meenu, S.P., Shisong, M., Tessa, M.B.S., Kirk, C., Peter, H., Savitharamma, P.D.K. (2013): Novel positive regulatory role for the SPL6 transcription factor in the N TIR-NB-LRR receptor-mediated plant innate immunity. - PLoS Pathogens 9: e1003235.

[47] Park, W., Scheffler, B.E., Bauer, P.J., Campbell, B.T. (2010): Identification of the family of aquaporin genes and their expression in upland cotton (*Gossypium hirsutum* L.). - BMC Plant Biology 10: 142.

[48] Rao, G., Sui, J., Zeng, Y., He, C., Duan, A., Zhang, J. (2014): De novo transcriptome and small RNA analysis of two chinese willow cultivars reveals stress response genes in *Salix matsudana*. ed. Matty Knight. - PLoS ONE 9: e109122.

[49] Sakuma, Y., Liu, Q., Dubouzet, J.G., Abe, H., Shinozaki, K., Yamaguchi. (2002): DNA-binding specificity of the ERF/AP2 domain of *Arabidopsis* DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. - Biochemical and Biophysical Research Communications 290: 998–1009.

[50] Seki, M., Kamei, A.,Yamaguchi-Shinozaki, K., Shinozaki, K. (2003): Molecular responses to drought, salinity and frost: common and different paths for plant protection. - Current Opinion in Biotechnology 14: 194–99.

[51] Shanker, A.K., Maddaala, A., Kumar, M.A., Yadav, S.K., Maheswari, M. (2012): In silico targeted genome mining and comparative modelling reveals a putative protein similar to an *Arabidopsis* drought tolerance DNA binding transcription factor in chromosome 6 of *Sorghum bicolor* genome. - Interdisciplinary Sciences: Computational Life Sciences 4: 133–41.

[52] Sharma, M.K., Kumar, R., Solanke, A.U., Sharma, R., Tyagi, A.K., Sharma, A.K. (2010): Identification, phylogeny, and transcript profiling of ERF family genes during

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200

development and abiotic stress treatments in tomato. - Molecular Genetics and Genomics 284: 455–75.

[53] Shigyo, M., Hasebe, M., Ito, M. (2006): Molecular evolution of the AP2 subfamily. - Gene 366: 256–65.

[54] Song, X., Li Y., Hou X. (2013): Genome-wide analysis of the AP2/ERF transcription factor superfamily in chinese cabbage (*Brassica aapa* Ssp. Pekinensis). - BMC Genomics 14: 998–1009. http://www.biomedcentral.com/1471-2164/11/719.

[55] Verheye, W. (2010): Growth and production of maize: traditional low-input cultivation. - Encyclopedia of Life Support Systems 1.

[56] Wu, H., Lv, H., Li, L., Liu, J., Mu, S., Li, X., Gao, J. (2015): Genome-wide analysis of the AP2/ERF transcription factors family and the expression patterns of DREB genes in moso bamboo (Phyllostachys Edulis). - Plos One 10: e0126657.

[57] Kazugo, M., Liu, Q., Miura, S., Shinozaki, Y.K., Shinozaki, K. (1999): Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. - Nature Biotechnology 17: 287–91.

[58] Yamasaki, K., Kigawa T., Seki, M., Shinozak, K., Yokoyama, S. (2013: DNA-binding domains of plant-specific transcription factors: structure, function, and evolution. - Trends in Plant Science 18: 267–76.

[59] Yan, Yan, H.W., Hong, L., Zhou, Y.Q., Jiang, H.Y., Zhu, S.W., Fan, J . (2013): A genome-wide analysis of the erf gene family in sorghum. - Genetics and Molecular Research 12: 2038–55.

[60] Yao, M., Xu, Z.S., Huang, Y., Tian, C., Wang, F., Xiong, A.S. (2015): Genome - wide analysis of AP2 / ERF transcription factors in carrot (*Daucus carota* L.) reveals evolution and expression profiles under abiotic stress. - Molecular Genetics and Genomics 290: 2049-6.

[61] Zarei, A., Körbes, A.P., Younessi, P., Montiel, G., Champion, A. (2011): Two GCC boxes and AP2/ERF-domain transcription factor ORA59 in jasmonate/ethylene-mediated activation of the PDF1.2 promoter in *Arabidopsis*. - Plant Molecular Biology 75: 321–31.

[62] Zhang, Kori, N.K., Zhang, C.H., Shangguan, L.F., Ma, R. J., Sun, X., Tao, R. (2012): Genome-wide analysis of the AP2/ERF superfamily in peach (*Prunus persica*). - Genetics and molecular research 11: 4789–4809.

[63] Zhang, G., Chen, M., Chen, X., Xu, Z., Guan, S., Li, L. C. Ma, Y. (2008): Phylogeny, gene structures, and expression patterns of the ERF gene family in soybean (*Glycine max* L.).- Journal of Experimental Botany 59: 4095–4107.

[64] Zhang, J.Z., Creelman, R., Zhu, J. (2004): From laboratory to field. using information from *Arabidopsis* to engineer salt, cold, and drought tolerance in crops. - Plant Physiology 135(2): 615–21.

[65] Zhang, Z., Huang, R. (2010): Enhanced tolerance to freezing in tobacco and tomato overexpressing transcription factor TERF2/LeERF2 is modulated by ethylene biosynthesis. -Plant Molecular Biology 73: 241–49.

[66] Zheng, Z., Mosher, L.S., Fan, B., Klessig, D.F. Chen, Z. (2007): Functional analysis of *Arabidopsis* WRKY25 transcription factor in plant defense against *Pseudomonas syringae*. -BMC Plant Biology 7(1): 2.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1405_177200