

# Genome-Wide in Silico Analysis of Ethylene Responsive Factors Associated with Stress Responses in Wheat

**Hareem Aamer**

National University of Sciences and Technology

**Alvina Gul** (✉ [alvina\\_gul@yahoo.com](mailto:alvina_gul@yahoo.com))

National University of Sciences and Technology <https://orcid.org/0000-0002-7323-1905>

**Rabia Amir**

National University of Sciences and Technology

**Muhammad Qasim Hayat**

National University of Sciences and Technology

**Faiza Munir**

National University of Sciences and Technology

**Syeda Marriam Bakhtiar**

Capital University of Science and Technology

**Mustafeez Babar**

Shifa Tameer-e-Millat University

**Zoya Khalid**

FAST-NU: National University of Computer and Emerging Sciences

---

## Research article

**Keywords:** Triticum aestivum, phylogenetic analysis. Motif analysis, gene ontology, sub-cellular localization, network analysis

**DOI:** <https://doi.org/10.21203/rs.3.rs-70071/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Transcription factors have largely been involved in the qualitative and quantitative regulation of gene expression. Among the diverse transcription factor families, AP2 (Apetala2) /ERF (Ethylene responsive factors superfamily) is a well-known plant family of proteins for controlling various and diverse stress responses in plants. Ethylene responsive factors (ERF) are known to be vital for generating various stress responses.

**Results:** A total of 181 *TaERF* proteins were identified in *Triticum aestivum*. Phylogenetic analysis, conserved motif analysis, chromosomal localization, gene structure analysis and multiple sequence alignment of ERF genes along with protein-protein interaction network analysis was performed. The phylogenetic analysis led to classification of these ERF members into 10 major groups and provided evolutionary relationship among the groups on the basis of similar protein motifs like which include WLG and 15 others. Further analysis revealed exon number, molecular weight, isoelectric point, and length of proteins. Gene ontology analysis concluded that *TaERF* were involved in responses to many Stressors, revealing their possible function. Sub cellular localization revealed that most of the proteins were confined to nucleus, cytoplasm and mitochondria. Network analysis provided insight of proteins involved in generating responses to stresses based on topological analysis and betweenness centrality of the network based on network theory.

**Conclusions:** Genome-wide identification, gene structure, phylogeny and network analysis of *TaERF* proteins in *Triticum aestivum* supplied a solid theoretical foundation for the functional study of *TaERF* family. The bioinformatics analyses of 181 *TaERF* were implemented systematically revealing proteins characteristics and structural information.

## Background

Abiotic stresses and biotic stresses affect many economical crops and reduce their yield and trigger stress responses [1]. Regulation of stresses occur at cellular, biochemical, physiological and molecular level. A number of stress responsive genes and proteins associated with signaling at molecular level have been identified and characterized [2]. Transcription factors are DNA binding proteins that initiate transcription and bind themselves to cis-regulatory elements. The functional and evolutionary similarities among different TFs have led to their classification into various families including APETALA2/ethylene response factor (AP2/ERF), AP2/ERF, WRKY, FAR1, DREB etc [3]. AP2/ERF transcription factor family play a crucial role in expression of genes, related to stress responses, reproduction, defense, and hormone secretion [4]. Over the past several years, there has been an increasing interest in the study of functional and structural characteristics of AP2/ERF. AP2/ERF superfamily emerged by horizontal transfer from bacteria/viruses to plants [5]. First discovered in *Arabidopsis thaliana*, the proteins comprises of 60 to 70 amino acids [6]. These genes were cloned or replicated in tobacco plants initially [7]. Currently, AP2/ERF gene family have been recognized in several species including wheat, barley, foxtail millet, sorghum, maize, barley, soya bean, grape, poplar, moso bamboo and among most of the members of the grass

family [8–11]. AP2, ERF and RAV based on the domain observed among these proteins. AP2 family have two AP2/ERF domains and is majorly involved in the plant developmental process like leaf formation and development of, flowers, embryos, ovules and fruits [12]. It also consists of two sub families AP2 and ANT [13]. ERF family has only one AP2/ERF domain and is involved in ethylene signal transduction. [14, 15], pathogen-stimuli and gene expression [16]. Recently, ERF genes have been studied in maize [17]. In RAV family, two structural domains are present on AP2/ERF and one B3 domain also present in other transcription factors. [18] Moreover, RAV show hormonal responses to ethylene and brassinosteroid [19].

ERF family is further split up into two subfamilies: the CBF (C-repeat)/DREB (dehydration response element binding) and ERF (Ethylene responsive factors) [20]. ERF subfamilies bind to the GCC box (AGCCGCC) [21], where G2, G5 and C7 are the core and crucial residues [21–24]. The 3-D analysis of AP2/ERF protein domain depict 3 anti-parallel beta-sheet and one alpha-helix [22–24]. On the other hand, ERF transcription factor proteins are involved in various signaling cascades involving complex protein-protein interactions (PPIs) forming protein networks driving various signaling pathways. PPIs are vital for a variety of cellular processes, thereby, making the apprehension of PPIs a crucial factor for perception of cell physiology in normal states and diseased states. [25–30]. Protein-protein interaction networks (PPIN) are mathematical depictions of the physical connections between various proteins within the cell. Globally, wheat demand is increasing with increase in world population and food demands [31]. Biotic and abiotic stresses affect wheat production to a greater extent. Many TF have been studied in wheat that respond to pathogen related attack and tolerating freezing temperatures [31–32]. To date, there is no report that represent detailed analysis of structure and function prediction of ERF proteins in wheat using in silico approaches. Genome-wide analysis of a family is relatively effective approach to classify and differentiation of plant gene or proteins functions, facilitating the study of the phylogeny of genes and genomes.

Though, AP2/ERF proteins have been thoroughly studied in various plant families, the current research provides detailed structural analysis and protein interaction network of ERF proteins involved in stress regulation. Present study is mainly focused on genome-wide analysis of ERF subfamily and protein-protein interaction network analysis in *Triticum aestivum*. Phylogenetic evolutionary studies, gene structure analysis, protein-protein interaction network analysis, gene ontology has been incorporated. This prediction-based study would help in providing an insight for future strategies and research for functional analysis and to design wet lab-based experiments.

## Results

### Bioinformatics analysis of ethylene responsive factors

ERF genes comparative analysis in different species was done showing *Triticum aestivum*, sorghum have almost similar number of ERF gene e.g. 181, and 172 respectively. While *Arabidopsis*, *Oryza sativa* and *Hordeum vulgare* have fewer number than *Triticum aestivum* (Table S1). The individual genes of *Triticum aestivum* are mentioned in (Table S2), with their predicted characteristics, including intron number, amino

acid (a.a) sequence length, molecular weight and isoelectric point. The length of proteins ranges from minimum 48 amino acids to maximum 1409 amino acids and PI ranged from 4.4 to 12.6 mentioned in (Table S2).

## MSA (Multiple sequence alignment)

To perceive and identify residues that are conserved in ERF domain, multiple sequence alignment of these 181 ERF genes was done using CLC work bench software package. With setting parameters: start gap: 10.0, cost setting as: 0.1, end gap cost: free. These parameters provide the insight of how the sequences will align which characters are to be considered as gap length of the gap. The ERF proteins sequences after Alignment were used to view phylogenetic tree in CLC viewer 20.0.3 and constructed in clustal w. Most sequences show similarities and are highly conserved across the family. The motifs such as WLG are conserved and present in most of protein sequences of wheat ERF's. Residues present in between are more or highly conserved collate to C-terminal and N-terminal residues as shown in (Fig. 1).

### Phylogenetic relationship between the ERF genes in *Triticum aestivum*

Phylogenetic analysis of 181 ERF proteins of *Triticum aestivum* was done with *Arabidopsis thaliana*, *Oryza sativa*, *hordeum vulgare*, *zea mays*. On the basis of phylogenetic analysis, ERF protein were seen to be closely related with *zea mays* and *hordeum vulgare*. NJ method with bootstrap replicates value 1000 was used. [Fig. 2].

## Subcellular localization and gene analysis of *Triticum aestivum* ERF

In order to obtain insight of ERFs structure, intron was analyzed by obtaining number of exons in each ERF. Most of the ERFs have intron and many of them have 1 or 2 introns. The number of introns were predicted by GSDS tool. Traes\_5BL\_7F0FD1538.2 protein has a greater number of exons (9). In subcellular localization analysis, most of ERF proteins were in nucleus. Plant-mPLOC evaluation, suggested that ERF genes were present in nucleus, cytoplasm, mitochondria. One protein was present in mitochondria and, Wolf psort predicted ERF genes located in multiple locations (Chloroplast, Nucleus, cytoplasm, Mitochondria, vacuole, cyto skeleton) as shown in (Table S3) and (Table S4).

## Conserved motif Analysis in ERF family

Conserved motif Analysis was done by MEME motif discovery the software depicted that there was ERF motif and many other motifs. Some ERF protein sequences also showed EAR-like motifs like other various plant species include *Arabidopsis thaliana*, rice plant and *Zea mays*. WLG motif was conserved among most of the sequences. MEME software for motif discovery inputs multiple protein sequences and number of conserved motifs as well as all input sequences motifs 1,2,3,4,5,6,8,11 are conserved across all ERF family with probability values as shown in [Fig. 2]

## Scanning PPI'S

PPIs were obtained, from STRING database the list of proteins were searched in the databases from which 123 out of 181 proteins showed interaction among them with a few other proteins of other transcription factors family. Every linkage gives data about hubs edges connections and what hubs and linkages demonstrate to and which color depict to which linkage.

## **PPIs network construction and with draw sub network from the network**

For the construction of PPIs network PPI interactions were taken from string database. The candidate genes were browsed into the STRING database and a file was obtained consisting all genes PPI interactions. A network was obtained by using drawing feature of PAJEK software. Topological analysis was done containing properties of hubs including degree (k), betweenness centrality (BC) and closeness centrality (CC) were used to assess the hubs in a system; particularly k and BC are two crucial boundaries in the network theory further more backbone network was constructed made up of high BC proteins and the links between them. 10% of the total node set was set as high BC to get most number of nodes as the backbone in order to study highly connected proteins in the analysis network. High BC value proteins were the candidates for hub proteins involved in various pathways as shown in (Fig. 4) and shown in (Table S5).

## **GO Analysis**

Gene ontology analysis predicted ERF proteins involvement in various biological processes, cellular processes, and molecular processes. The analysis depicted that ERF proteins were present in different stresses and their responses and involved in different regulations. Stress responses in which ERFs were involved include water stress, hormones, cold and salt stress. In different molecular processes, ERF proteins were predicted to possess DNA binding activity as shown in (Fig. 5).

## **Discussion**

Transcription factors presume a vital role in controlling and the adjusting responses of plants to different inside or outside signals [26]. They regulate the downstream genes in stress signal transduction pathways by means of initiation and suppression of genes on exposure to stresses. Plant genomes contain multiple transcription factors like AP2/ERF [27]. It has been anticipated that the Arabidopsis genome codes for 1533 transcription factors, involving over 5.9% of its absolute anticipated genes (Rao et al., 2014). As per, 181 AP2/ERF genes or proteins were recognized in ERF subfamily having one Apetala2-like domain or region. All through the most recent decade, a colossal measure of exploration has been coordinated that demonstrated overexpression of ERF family genes builds the resistance of plants to stresses as studied in [25].and Shinozaki, 2012) in chillies [25–30]. Therefore, it is one of the most significant challenges to distinguish the genes or proteins which provide resistant against abiotic factor. The phylogeny, structural analysis, GO analysis, motif analysis in various plants incorporating genome-wide investigation in model plant, Arabidopsis, rice [25]. And sorghum [26]. While few ERF genes

in wheat were additionally concentrated in pathogen attack. With the completion of wheat genome sequencing ventures, phylogenetic examination of TaERF proteins or genes would be useful to explore general function of ERF genes or proteins and the developmental processes of the AP2/ERF in plants. It was found that TaERF has 181 ERF genes which are more noteworthy than Arabidopsis, sorghum, peach and moso bamboo, while not as much as Chinese cabbage and Carrot.

A comparative example of conserved residues was accounted for in sorghum ERF genes [27].and in cucumber. [28]. these conserved residues have critical impact on the function of AP2-domain and furthermore help for single change in protein. Comparative analysis of amino acid residues among various plants have demonstrated that they are most conserved in the AP2/ERF superfamily were available in different species, including Arabidopsis, rice, maize and different plants like WLG and EAR [25]. Are additionally seen in TaERF proteins in which WLG is the most important. Aside from motifs in Arabidopsis and rice, TaERF assume significant role in evolutionary relationships. A significant number of exons were found in TaERF proteins. 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 representation showed that the presence of exons in ERF genes is also a feature of TaERF. Resulting in diverting the biological, cellular and chemical processes including abiotic stress responses like drought, cold, temperature and salinity. Our findings are similar to previous study of AP2/ERF genes in other species, like ZmERF. The molecular process of TaERF showed Cellular localization showed most genes are located in nucleus, few in the mitochondria and other organelles. This means that some TaERF have multiple locations in the cell. This characteristic of AP2/ERF genes was reported in ZmERF.

Apart from these findings the protein protein interaction network was analyzed and it showed various proteins are physically connected with each other having high BC values and shortest path length which means these proteins having high BC values are hub proteins playing a role in regulating the whole network these are the directly linked proteins one protein can link to several proteins at a time which means that protein has high BC value.

To our knowledge, this is the first study in Wheat *Triticum aestivum* for genome-wide analysis of ERF genes. This research may also contribute functional gene resources for genetic future engineering approaches for producing plants resistant to various biotic and abiotic stresses. This research will also provide core proteins involved in a network pathway that could be used for further investigations in future. ERF's with high BC value are involved in various pathways and are involved in various functions. The structural Analysis is linked to the network analysis as it represents the deeper insight of each ERF which help us identify sites which are important, and which are an apart of various processes. Phylogenetic Analysis reveal the ancestral groups containing similar ERF's such as arabidopsis and other grasses.

## Conclusion

As wheat genome has been sequenced with a lot of information about the genome still there is a lot of information to be interpreted from the genome which is ambiguous and needed to be studied further such as functions of certain genes that are unidentified and need further analysis there are a lot of wheat transcription factors that are not well analyzed and have unclear information that needs to be studied. This study aims to provide a detailed structural and PIN analysis on wheat ERF's based on prediction using bioinformatics and computational biology. This research may also contribute functional gene resources for genetic future engineering approaches for producing plants resistant to certain biotic and abiotic stresses. This research will also provide core proteins involved in a network pathway or signaling pathways initiating stress responses that could be used for further investigations in future however, further research needs to be done for better understanding.

## Methods

### Identification of Ta ERF proteins by computational Analysis

AP2/ERF superfamily has thoroughly been studied in *Triticum aestivum* which led to the identification of 183 ERF subfamily proteins. The 183 ERF genes identified and reported in an earlier study were considered for the study. Data of these genes were obtained from different databases including Plant Transcription Database (<http://plantfdb.cbi.pku.edu.cn/>), Phytozome (<http://phytozome.jgi.doe.gov/>). As a result, sequences of 183 ERF genes, 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>).

### Phylogenetic-Domain- MSA analysis of ERFs

To observe conserved residues in ERF domain, multiple sequence alignment of these 181 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 20.0.3 (Knudsen et al., 2011) by employing Neighbor Joining method, Distance measure: Jukes-Cantor and Bootstrap: 1000 Replicates In order to ensure high level of validity, 1000 bootstrap replicates were performed. This is an approach used for determining and calculating evolutionary and phylogenetic relationships among species.

### ERF protein motif analysis

Motif analysis was performed using MEME online software (Bailey and Elkan, 1994; Bailey et al., 2006). The descriptors for data analysis included; number of different motifs: 20, minimum motif width: 15, maximum motif width: 60 amino acids, distribution of the motif occurrence: zero or one per sequences. These parameters were incorporated according to the length of the protein and number of motifs under study to narrow down the desired outcomes.

### Scanning PPI's and construction of Protein- Interaction network

PPIs were taken, from STRING database (<https://string-db.org/cgi/network.pl>) for the inspection of protein–protein interactions. A network was constructed that consist of the seed proteins and also their close neighbors and the linkage between all proteins. Pajek program was used for the analysis of network linkages, construction and visualization of huge networks.

## Topological graph theory analysis

Properties of graph theory were acquired to evaluate the nodes or center points in a network; most important of them are  $k$  which is degree of the network and BC which is the centrality of the network or the shortest paths among nodes are essential and necessary in the network theory. For topological center of the network worldwide topological estimations of systems incorporate degree, mean, path length and diameter used to define network. The proteins with high BC values were the intensely seed or hub centers, and the connections between them make up a backbone. The most direct route between the candidates were calculated and evaluated by Pajek software (<http://mrvar.fdv.uni-lj.si/pajek/>). The sub network construction was done by shortest or longest path calculation with nodes having shortest path be the hub proteins we detected the direct routes between selected proteins which aided in obtaining the proteins having multiple interactions and centrality determining to be the centre or highly involved proteins of the network.

## GO Analysis and sub cellular localization

The gene ontology analysis of ERF protein sequences of wheat was performed using BLAST GO with default settings. The number of exons using GSDS tool were predicted. Sub-cellular localization was performed by two web-based tools WoLF PSROST and Plant-mPLOC. GO analysis of wheat ERF provided information of the biological process in which ERFs are involved, molecular function they perform in the cell. Information about cellular localization where they are located in cell which provides the insight of the organelle in which they are situated and helps in predicting their possible function in the cell that could contribute against different stress responses.

## Abbreviations

TFs: transcription factors; TaERF: Triticum aestivum Ethylene responsive factors; CBR: C-element binding factor; PPIN: Protein protein interaction network. GO: gene ontology NJ: Neighbor joining

## Declarations

### Acknowledgements

Not applicable.

### Authors' contributions

HA designed the research performed the analysis and other authors did the proof reading and interpretation. AG supervised the whole project. RA, MQH and FM reviewed the research. SMB provided

the overall guidelines for analysis. MB and ZK helped in analysis.

All authors have read and approved the manuscript.

### **Availability of data and materials**

All data and materials used in this study are attached.

### **Ethics approval and consent to participate**

Not applicable.

### **Consent for publication**

Not applicable.

### **Competing interests**

The authors declare that they have no competing interests.

### **Author details**

Hareem Aamer, Alvina Gul, Rabia Amir, Muhammad Qasim Hayat, Faiza Munir: Department of Plant Biotechnology, Atta-ur-Rahman School of Applied Biosciences, National University of Sciences and Technology, Islamabad, Pakistan

Syeda Marriam Bakhtiar, Department of Bioinformatics and Biosciences, Capital University of Science and Technology, Islamabad, Pakistan

Mustafeez Babar, Department of Biosciences, Shifa Tameer-e-Millat University, Islamabad, Pakistan

Zoya Khalid, Department of Computer Science, NUCES-FAST, Islamabad, Pakistan

All authors have read and approved the manuscript.

## **References**

1. Boyer JS. Plant Productivity and Environment. *Science*. 1982;218:443–8.
2. Zhuang J, Cai B, Peng RH, Zhu B, Jin XF, Xue Y, et al. Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. *Biochem Biophys Res Commun*. 2008;371:468–74.
3. Pabo CO, Sauer RT. Transcription factors: structural families and principles of DNA recognition. *Annual review of biochemistry*. 1992;61:1053–95.
4. Manuela Hinz IW, Wilson J, Yang, Katharina Buersten binder, D, Llewellyn ES Dennis, Margret Sauter, Rudy Dolferus. *Plant Physiology* Jun 2010, 153 (2) 757–772.

5. Xu Z-S, Chen M, Li L-C, Ma Y-Z. Functions and Application of the AP2/ERF Transcription Factor Family in Crop Improvement F. *J Integr Plant Biol.* 2011;53:570–85.
6. Jofuku KD, den Boer BG, Van Montagu M, Okamoto JK. Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *Plant Cell.* 1994;6(9):1211–25.
7. Ohme-Takagi M, Shinshi H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell.* 1995;7(2):173–82.
8. Zhuang J, Cai B, Peng RH, Zhu B, Jin XF, Xue Y, et al. Genome-wide analysis of the AP2/ERF gene family in *Populus trichocarpa*. *Biochem Biophys Res Commun.* 2008;371:468–74. doi:10.1016/j.bbrc.2008.04.087.
9. Song X, Li Y, Hou X. Genome-wide analysis of the AP2/ERF transcription factor superfamily in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). *BMC Genom.* 2013;14:573.
10. Wu H, Lv H, Li L, Liu J, Mu S, Li X, et al. Genome-wide analysis of the AP2/ERF transcription factors family and the expression patterns of DREB genes in Moso bamboo (*Phyllostachys edulis*). *PLoS ONE.* 2015;10:e0126657.
11. Lata C, Mishra AK, Muthamilarasan M, Bonthala VS, Khan Y, Prasad M. Genome-wide investigation and expression profiling of AP2/ERF transcription factor superfamily in foxtail millet (*Setaria italica* L.). *PLoS ONE.* 2014;9:e113092. doi:10.1371/journal.pone.0113092.
12. Xu Z-S, Chen M, Li L-C, Ma Y-Z. Functions and Application of the AP2/ERF Transcription Factor Family in Crop Improvement F. *J Integr Plant Biol.* 2011;53:570–85.
13. Ohme-Takagi M, Shinshi H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell.* 1995;7(2):173–82.
14. Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. OsDREB genes in rice, *Oryza sativa* L., encode transcription activators that function in drought-, high-salt- and cold-responsive gene expression. *The Plant journal: for cell molecular biology.* 2003;33(4):751–63.
15. Zarei M, Azizi M, Bashir-Sadr Z. Evaluation of physicochemical characteristics of pomegranate (*Punica granatum* L.) fruit during ripening. *Fruits.* 2011;66:121–9. 10.1051/fruits/2011021.
16. Hussain A, Sultan S, Hussain S, Ahmed I, Mustafa, Ghulam, Shahid, Muhammad. Genome-wide analysis of ethylene responsive factor in maize: An in silico approach. *Appl Ecol Environ Res.* 2016;14:177–200. 10.15666/aeer/1405\_177200.
17. Kagaya Y, Ohmiya K, Hattori T. 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.* 1999;27(2):470–8.
18. Chitnis VR, Gao F, Yao Z, Jordan MC, Park S, et al. After-Ripening Induced Transcriptional Changes of Hormonal Genes in Wheat Seeds: The Cases of Brassinosteroids, Ethylene, Cytokinin and Salicylic Acid. *PLOS ONE.* 2014;9(1):e87543.
19. 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(3), 998–1009.
20. Ohme-Takagi M, Shinshi H. Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. *Plant Cell*. 1995;7(2):173–82.
  21. Hao C-M, Redha R, Morrow J, Breyer M. (2002). Peroxisome proliferator-activated receptor  $\delta$  activation promotes cell survival following hypertonic stress. *The Journal of biological chemistry*. 277.21341-5. 10.1074/jbc.M200695200.
  22. Maryan KE, Lahiji HS, Farrokhi N. Hassan Hasani Komeleh, Analysis of Brassica napus dehydrins and their Co Expression regulatory networks in relation to cold stress, *Gene Expression Patterns*, Volume 31,2019,Pages 7–17,ISSN 1567-133X,.
  23. Barabasi AL, Albert R. Emergence of scaling in random networks. *Science*. 1999;286(5439):509–12.
  24. Deane CM, et al. Protein interactions: two methods for assessment of the reliability of high throughput observations. *Mol Cell Proteomics*. 2002;1(5):349–56.
  25. Yu H, et al. The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput Biol*. 2007;3(4):e59.
  26. Ideker T, et al. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*. 2002;18(Suppl 1):S233-40.
  27. Huang daW, et al. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1–13.
  28. Girvan M, Newman ME. Community structure in social and biological networks. *Proc Natl Acad Sci U S A*. 2002;99(12):7821–6.
  29. Xiuliang Zhu L, Qi, Xin Liu S, Cai H, Xu R, Huang J, Li X, Wei. Zengyan Zhang *Plant Physiology Mar*. 2014;164(3):1499–514.
  30. Verheye W. Growth and production of maize: traditional low-input cultivation. In: Verheye WH, editor. *Land use, land cover and soil sciences*. Oxford: UNESCO-EOL; 2010.
  31. Xiuliang Zhu, Lin Qi, Xin Liu, Shibin Cai, Huijun Xu, Rongfeng Huang, Jiarui Li, Xuening Wei, Zengyan Zhang *Plant Physiology Mar* 2014, 164 (3) 1499–1514.
  32. Verheye, W. (2010). Growth and production of maize : traditional low-input cultivation. In W. H. Verheye (Ed.), *Land use, land cover and soil sciences*. Oxford, UK: UNESCO-EOL

## Figures



## Figure 1

Multiple sequence alignment of TaERF performed by protein sequences using CLC workbench software package depicting a portion of alignment.



## Figure 2

The phylogenetic analysis of TaERF proteins; NJ tree was constructed from protein sequences of TaERF using Clustral W with 1000 bootstrap copies.



## Figure 2

The phylogenetic analysis of TaERF proteins; NJ tree was constructed from protein sequences of TaERF using Clustral W with 1000 bootstrap copies.

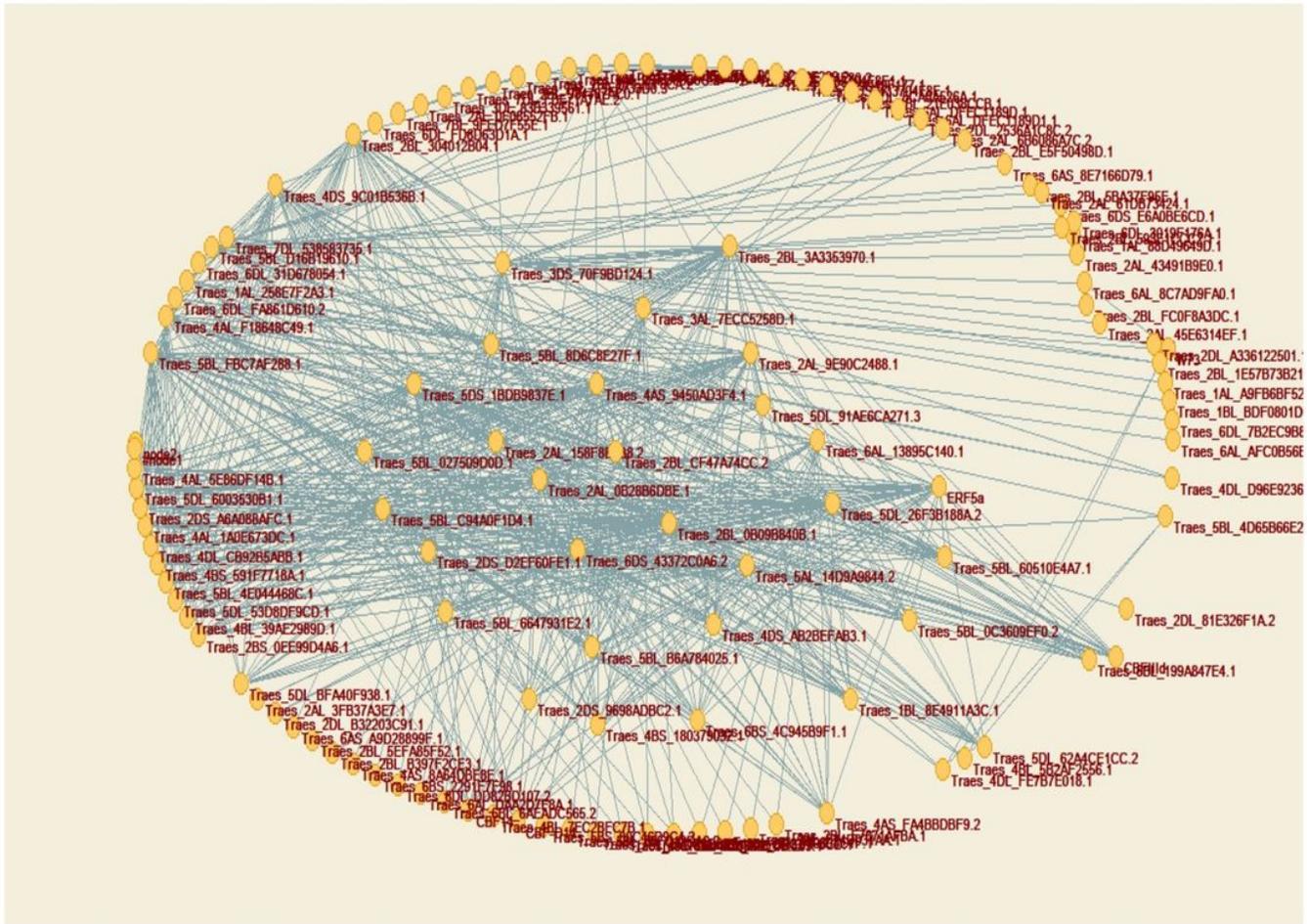
[DISCOVERED MOTIFS](#) | [MOTIF LOCATIONS](#) | [INPUTS & SETTINGS](#) | [PROGRAM INFORMATION](#) | [RESULTS IN TEXT FORMAT](#) | [RESULTS IN XML FORMAT](#)

DISCOVERED MOTIFS		E-value	Sites	Width	More	Submit/Download
1.		2.1e-2529	161	26		
2.		6.5e-2131	134	23		
3.		8.3e-498	160	11		
4.		3.3e-403	14	50		
5.		3.0e-254	26	29		
6.		1.1e-200	21	21		
7.		3.2e-188	9	50		
8.		2.1e-126	9	29		
9.		8.2e-122	6	42		
10.		1.4e-107	11	21		
11.		9.2e-101	16	29		
12.		2.2e-097	4	50		
13.		1.2e-084	4	50		
14.		3.3e-092	5	50		
15.		2.1e-073	13	21		

## Figure 3

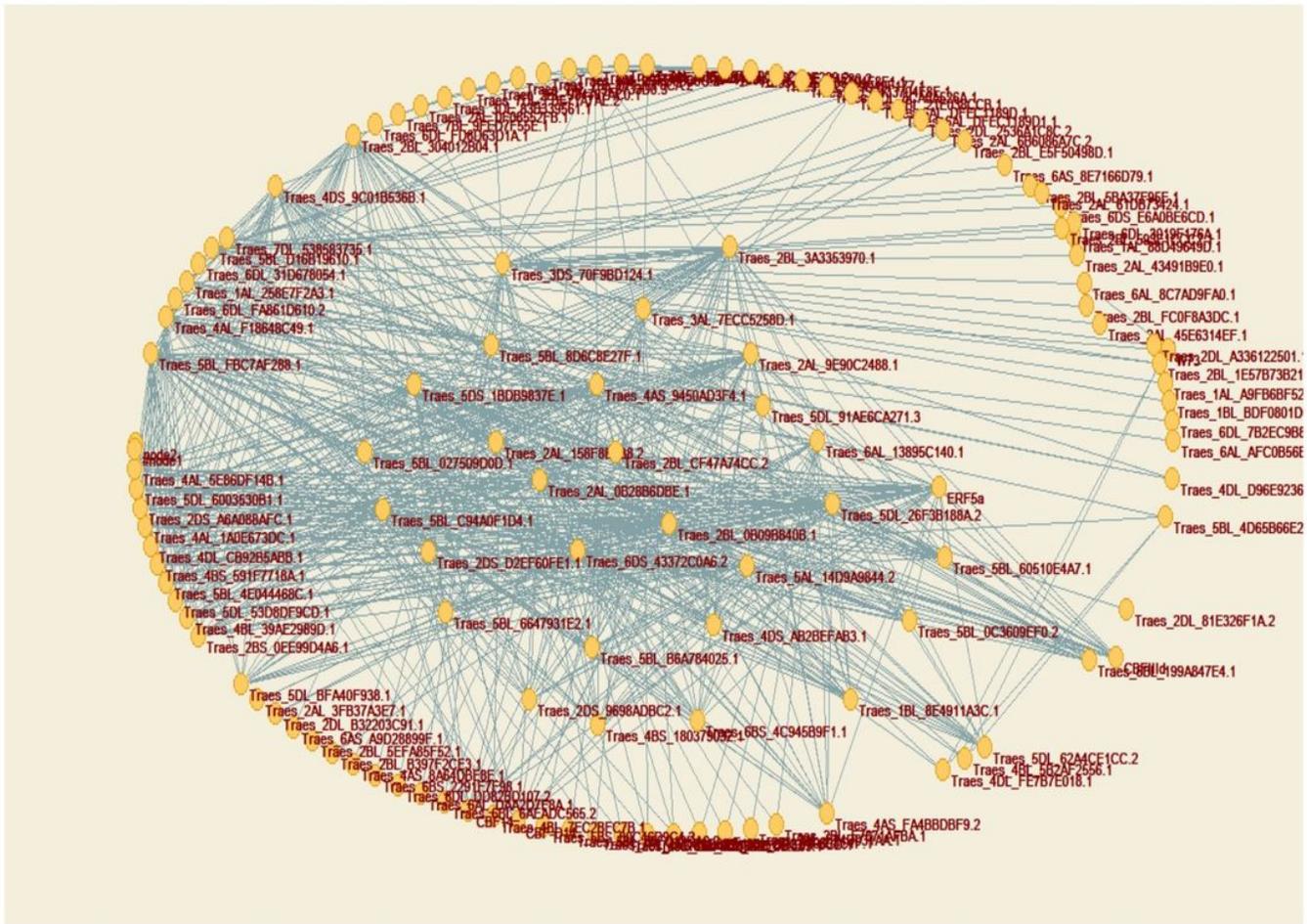
Conserved motifs of TaERF protein; the motif in the TaERF protein was identified by using Multiple Em for Motif Elicitation (MEME). Fifteen conserved motifs were identified and are displayed in different colors.





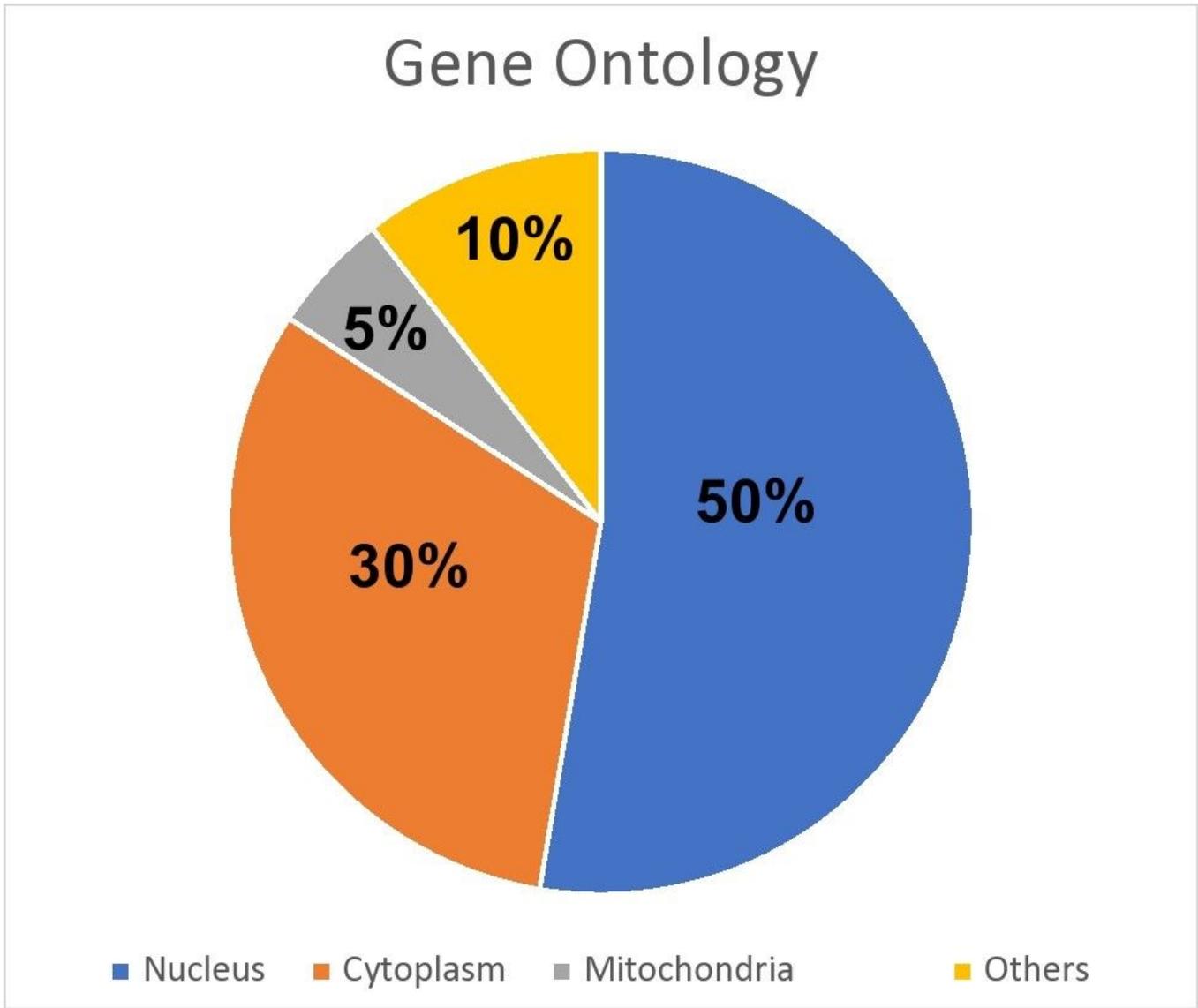
**Figure 4**

Network construction of *Triticum aestivum* ERFs through protein-protein interactions using pajek software.



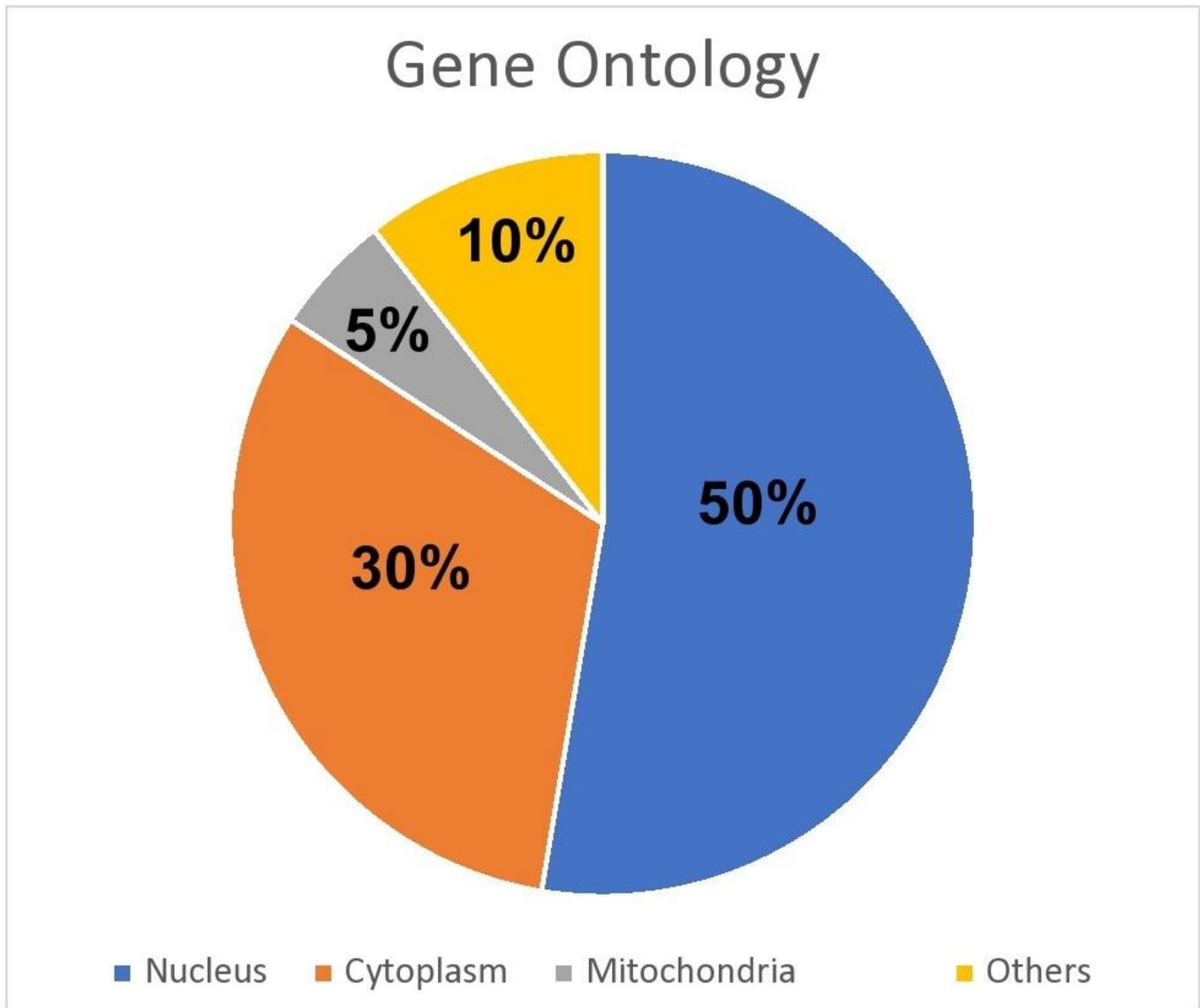
**Figure 4**

Network construction of *Triticum aestivum* ERFs through protein-protein interactions using pajek software.



**Figure 5**

The percentage graph of TaERF proteins sub-cellular localization.



**Figure 5**

The percentage graph of TaERF proteins sub-cellular localization.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [TableS1.pdf](#)
- [TableS1.pdf](#)
- [TableS2.pdf](#)
- [TableS2.pdf](#)
- [TableS3.pdf](#)

- [TableS3.pdf](#)
- [TableS4.pdf](#)
- [TableS4.pdf](#)
- [TableS5.pdf](#)
- [TableS5.pdf](#)