NARDI FUNDULEA, ROMANIA
www.incda-fundulea.ro

ROMANIAN AGRICULTURAL RESEARCH, NO. 39, 2022
First Online: January, 2022. DII 2067-5720 RAR 2022-103

PHYLOGENETIC AND PHYSICOCHEMICAL STUDY OF CBF GENE IN DIFFERENT PLANT SPECIES

Iman Yousefi Javan[*] and Masoud Alipanah

Department of Plant Production, Faculty of Agriculture, University of Torbat Heydarieh, Iran
[*]Corresponding author. E-mail: I.Javan@torbath.ac.ir

ABSTRACT

Temperature changes and environmental stresses are the most important non-living factors in disturbing the favorable conditions that cause metabolic disorders in plant cells. In this regard, CBF systems have the most important role in tolerating cold stress in plants. It is very important to study and compare the sequence of CBF gene and determine the phylogenetic and evolutionary relationships of this gene in different plants using bioinformatics tools. In this study, by examining the amino acid sequence of protein expressed from CBF gene in durum wheat (tetraploid), from Plantcare, NCBI and EXPASY databases and also to investigate the relationships of this gene in forty different plant species, their homology from in terms of protein similarity with MEGA5, UPGMA and Bioedit software was determined. Bioinformatics analysis showed that the nucleotide length of this gene among the forty plants studied above varied from 453 nucleotides for pine to 759 nucleotides for rice. It also showed that the length of this gene in hexaploid wheat plant contains 214 amino acids. The results of analysis of variance showed high diversity between plants. The evolutionary matrix was calculated based on Poisson correlation and showed that there is the highest distance (4.324) and the lowest distance (0.007) between the protein sequences of CBF gene in plants. Based on the results of cluster analysis by UPGMA method, the protein sequence of this gene was divided into three groups. AP2 and ABRE regions were observed in the initial sequence of CBF gene and also the presence of CAAT box in most plants. The existence of genetic diversity of CBF gene in different plants confirms that the differences in gene expression and differences in protein form are due to environmental effects, which in the long run have epigenetic effects and genetic modification of that gene in that plant.

Keywords:CBF, genetic diversity, cluster analysis, AP2, protein spatial shape.

INTRODUCTION

Non-living stresses are one of the main causes of reduced crop yield worldwide, among which low temperature or salinity is the most important environmental factor and leads to the destruction of agricultural land and also reduced crop yield (Vaidyanathan et al., 2003). Sub-zero temperatures, especially sudden cold in spring and autumn, cause the buds to die. In fact, low temperature is a limiting factor for plant growth and distribution and affects the quality and quantity of plant production. Coping with low temperatures forces plants to respond to various stress mechanisms. And therefore show different degrees of tolerance. This great difference between plants in the field of tolerance to different temperatures stems from the wide response at the levels of cell walls, cell membranes, organelles, micromolecules, macromolecules, and finally the different expression of related genes.
Measuring the effect of various factors on tolerance to abiotic stresses is done through several methods, including measuring gene transcription levels (Loggini et al., 1999; Xin and Browse, 2000; Esfandiari et al., 2007; Xin et al., 2007; Maurya and Bhalerao, 2017). In cold climate, due to the amount of rainfall and the potential of agricultural soils in the region, the best way to deal with damage caused by cold and frost, modification of high-yielding genotypes and resistant to cold stress can be. The mechanisms of frost tolerance are not yet fully understood. However, effective mechanisms of frost tolerance may prevent the denaturation of proteins, the deposition of molecules, and the reduction of physical damage caused by the accumulation of ice in intercellular spaces. It also produces active

Received 18 September 2021; accepted 25 January 2022.

oxidant species such as superoxide (O) and hydroxyl radical (OH) in chloroplasts and mitochondria, which cause lipid peroxidation, protein denaturation, and DNA mutations. Together, these impairments lead to metabolic disorders that play a vital role and biomolecular role in cellular metabolism (Prerostova et al., 2020). Destructive biochemical pathways are activated by increasing endonuclease activity. Another important ecological stress in this regard is osmotic stress due to salinity. This reduces water absorption and determines gas exchange in the plant (Esfandiari et al., 2007). Ionic and osmotic stress due to salinity leads to oxidative stress in the plant and the cells are slightly reduced due to increased production of reactive oxygen species ROS. In this case, the processes of photosynthesis are disrupted (Eltayeb et al., 2007). Tolerance of different plant genotypes to cold stress is different due to their morphological or physiological characteristics (Filippi et al., 2007). The main goal of cold-related research has been to find genes that to be responsible for inducing cold and frost resistance of plants. Most genes responsible for cold resistance control the biochemical and physiological activities necessary for plant growth and development at low temperatures (Esfandiari et al., 2007).

Although it is possible to obtain cold-resistant plant cultivars by breeding methods under cold stress conditions, the probability of success in this field using conventional classical methods of plant breeding is low. The use of new technologies, especially plant biotechnology, to introduce and improve the nutritional status of 8 billion people in 2025 is an important step in introducing cold-resistant cultivars and can be successful. The use of bioinformatics methods to study cold-resistant genes, QTL study related to biotechnology projects is the best and most reliable method for the production of cold-resistant cultivars, which can control the chromosomal regions of the trait and the genetic effects of these regions. Option to find cold-resistant cultivars by assessing plant survival in the field by researchers has been done (Eltayeb et al., 2007). The proteins that play a role in plant adaptation and cold tolerance when exposed to cold are plant transcription factor DREB or CBF. Numerous mechanisms respond to freezing stress through physiological, biochemical, and molecular cellular processes, adapting to environmental conditions and thus showing different amounts of tolerance. This difference between plants stems from the large response at the cell wall and membrane surfaces, organelles, micromolecules, and macromolecules, and ultimately leads to different expression of related genes (Filippi et al., 2007). In Arabidopsis, many studies have shown that the CBF gene pathway is involved in adapting to cold and fighting colds. The family of CBF regulators plays a central role in stress tolerance in plants (Thomashow, 2001; Shi et al., 2017).

Gene CBF/DREB1 (C-repeat binding factor/dehydration reaction-responsive element-binding) encrypts a small family of transcription activators. To determine this role, a reverse genetic approach and a specific mutation were used. Also, when plants are briefly exposed to low temperatures, transcription of the CBF and DREB genes in those plants increases dramatically (Jaglo-Ottosen et al., 1998; Medina et al., 1999). Specifically, the CBF gene also adapts to the process of salinity stress and drought (Haake et al., 2002). The CBF gene actually encodes transcriptional activators that are members of the AP2 family and are DNA-binding proteins. These transcription factors are cold and drought dependent and are DNA regulatory elements (CRT responsive elements) (Baker et al., 1994; Stockinger et al., 1997; Shi et al., 2018). In general, CRT/DRE elements are present in COR promoters as well as in many other important genes resistant to cold stress to increase plant freezing tolerance (Stockinger et al., 1997; Haake et al., 2002). Some studies have shown that several mechanisms may contribute to increased tolerance to freezing, salinity, and drought. One of these important mechanisms is the synthesis of apparent protective polypeptides such as COR15a (Park et al., 2015; Zhao et al., 2016). Accumulation of this polypeptide

causes protective properties such as accumulation of sucrose, raffinose and proline. (Nanjo et al., 1999; Gilmour et al., 2000; Taji et al., 2002) The CBF gene is present in different plant species (Haake et al., 2002).

The DREB gene transcription factor, which is induced in response to salinity, dryness, and heat, increases plant resistance to stress by intensifying the expression of downstream genes. CBF gene expression, in turn, ensures proper induction of downstream genes to increase the expression of other cold-resistant genes (Novillo et al., 2004). Expression of CBF genes in transgenic plants increased frost resistance in adapted and non-adapted plants to cold, and those plants showed an increase in cold resistance (Jaglo et al., 2001; Hsieh et al., 2002). Regarding the CBF gene, it can be said that the orthologous DREB1/CBF genes have been identified in many crops such as broccoli, tomatoes, alfalfa, corn, rice, barley, atriplex, canola, eucalyptus, pistachios and wheat. DREB1/CBF is in the plant domain and CBF technology is also expected to improve stress tolerance, especially frost resistance of crops by controlling the expression of this regulatory system (Zhen and Ungerer, 2008). Today, bioinformatics sources and methods are used to analyze, interpret information, and identify relationships between data. One of these areas is molecular analysis, which involves sequencing sequences, searching databases, identifying images, and plotting evolutionary relationships and genomic comparisons. It is essential because of the importance of the CBF system as an effective system in ionic homeostasis. The sequence of this gene and its similarity in different plants were examined and compared. One of the objectives of this study is to study and compare the sequence of CBF gene and determine the phylogenetic and evolutionary relationships of this gene in different plants using relevant bioinformatics tools. For this purpose, in order to identify the molecular response of the plant to cold stress, an attempt was made to identify the CBF gene sequence in plants by bioinformatics methods. And then we tried to prove that the presence of CBF gene in the plant genome increases the freezing tolerance capacity of the plant.

MATERIAL AND METHODS

Grain Genes database (https://wheat.pw.usda.gov) was used to evaluate the function of CBF gene from tetraploid wheat (Turgidum wheat). CBF gene ID number, EF028778.1 was obtained from the DDBJ site (http://www.ddbj.nig.ac.jp).

Making RNA, cDNA and PCR reaction:

All RNA extraction steps were performed according to the instructions of Fermentase Company (add kit name for RNA isolation and for RT-PCR) for using Trizol. Then 1 µg of RNA, 0.6 picomol of OligodT and finally 8 µl of sterilized water were added to make cDNA. The resulting set was exposed to 80°C and then to -10°C. An amount of 0.5 µl of Rnasine, 4 µl of buffer, and 2 µl of dNTP were added at a concentration of 10 mM. The resulting mixture was added at 37 and 70°C after adding the enzyme to activate and deactivate the enzyme. Bio-Rad (add the name of PCR device was used for PCR reaction. 33 cycles were used with a temperature of 98°C for 45 s, a binding temperature of 58° for 40 s and a propagation temperature of 72° for 45 s (primers at 1 µl each, dNTP 3 µl, 10X buffer and $MgCl_2$, 5 µl, Taq polymerase enzyme, 1 µl, sterilized water 20 µl). Primers used to replicate cDNA are:

Forward Primer
CGCTCCTCTCTCAAGTGTCT

Reverse Primer
CACAATGAACGAGCACATA

After PCR reaction, the quality of the product on agarose gel at 1% concentration was evaluated.

Bioinformatics study of CBF gene

Sequences of this gene in different plants studied in this article were obtained from the NCBI Gene Bank (http://www.ncbi.nlm.nig.gov) plus nucleotide sequences from the SWISS-PROT database. In order to identify and study the CBF gene sequence and study the protein expressed from the above gene, primers were designed from the probable CBF gene sequence, due to the presence of fixed conserved sequences in specific regions of similar peptides such as the initial and terminal regions of the gene. Primer 3 software was used, also for more complete study of the above gene, information, registered databases were used. Then, for multiple sequencing studies, by blasting, nucleotide and protein sequences in NCBI database with other homologous sequences of CBF gene in other plants were identified and analyzed (Table 1). In this study, an attempt was made to select species that, despite the apparent sexual and interspecific differences, have many similarities in terms of the presence of this gene. Bioedit software was used to align the sequences and check the overlapping points, and MEGA V5 software was used to determine the evolutionary relationships and calculate the distance matrix. Cluster analysis was performed by UPGMA method with relevant software. After comparing the sequence of this gene in different plant species, the site http://bioinformatics.psb.ugent.be/webtools/plantcare/html was used to study the cis elements in the promoter regions of the CBF gene. Instruction, analysis and translation of the nucleotide sequence encoding the above protein, as well as physicochemical properties such as isoelectric point (IP), prediction of intramembrane regions, domain search and post-translational changes via http://expasy.org Analyzed. The three-dimensional structure of the protein of this gene was plotted using CATIA V3 software and the site http://swissmodel.expasy.org/interactive. Due to the need to study the interaction between protein amino acids, docking studies were performed using Molegro Virtual Docker (MVD) software.

Table 1. The names of the plants analysed in terms of amino acid sequences of proteins expressed from the CBF gene

No.	Latin name of the plant	The common name of the plant	Plant family name	Reference number in the gene bank
1	*Triticum aestivum C. Norstar*	Wheat bread	Gramineae	EF028778.1
2	*Triticum aestivum C. Jing411*	Wheat bread	Gramineae	AAX28963.1
3	*Triticum turgidum*	Wheat turgidum	Gramineae	CDN65459.1
4	*Triticum monococcum*	Wheat monococcum	Gramineae	EU076382.1
5	*Secal cereale*	Rye	Gramineae	AAL35759.1
6	*Lolium perenne*	Lolium	Gramineae	BAF36843.1
7	*Hippophae rhamnoides*	Oleaster	Elaeagnaceae	KU497732.1
8	*Arabis pumila*	Gilly flower	Brassicaceae	DQ207404.1
9	*Juglans regia*	Walnut	Juglandaceae	JX875914.1
10	*Thlaspi arvense*	Shepherd's bag	Brassicaceae	EU159411.1
11	*Sabal palmetto*	Palm leaf	Arecaceae	DQ497730.1
12	*Cocos nucifera*	Coconut	Arecaceae	DQ497739.1
13	*Dypsis lutescens*	Butterfly palm	Arecaceae	DQ497738.1
14	*Elaeis oleifera*	American oil palm	Arecaceae	DQ497734.1
15	*Trachycarpus fortunei*	Palm mill	Arecaceae	DQ497732.1
16	*Solanum melongena*	Eggplant	Solanaceae	KY780486.1
17	*Alyssum dasycarpum*	Qadomeh	Brassicaceae	JQ687133.1
18	*Oryza sativa*	Rice	Gramineae	AF243384.1
19	*Citrus sinensis*	Orange	Rutaceae	FJ861084.1
20	*Hordeum vulgare*	Barley	Gramineae	AF239616.1
21	*Brassica napus*	Canola	Brassicaceae	HM235815.1
22	*Raphanus sativus*	Radish	Brassicaceae	GQ866977.1
23	*Arabidopsis thaliana*	Arapidobsis	Brassicaceae	EF522964.1
24	*Capsella rubella*	Priest bag	Brassicaceae	XP_006284509.1
25	*Crocus alatavicus*	Wild saffron	Iridaceae	ASW18446.1

No.	Latin name of the plant	The common name of the plant	Plant family name	Reference number in the gene bank
26	*Chorispora bungeana*	Special smell night	Brassicaceae	AAY21899.2
27	*Camelina sativa*	False flax	Brassicaceae	XP_010448428.1
28	*Malcolmia scorpioides*	Desert nightshade	Brassicaceae	AGY36889.1
29	*Nicotiana tabacum*	Tobacco	Solanaceae	ABD65969.1
30	*Brassica juncea*	Chinese mustard	Brassicaceae	AAW79077.2
31	*Solanum lycopersicum*	Tomato	Solanaceae	NP_001234123.1
32	*Populus trichocarpa*	Poplar tree	Salicaceae	ABO48363.1
33	*Cucumis sativus*	Cucumber	Cucurbitaceae	ABG38530.1
34	*Pinus lambertiana*	Pine tree	Pinaceae	AEW08200.1
35	*Crocus sativus*	Saffron	Iridaceae	AWT62825.1
36	*Glycine max*	Soybean	Leguminosae	ABQ42206.1
37	*Fragaria vesca*	Strawberry	Rosaceae	ACN87752.1
38	*Vitis vinifera*	Grape	Vitaceae	AIL00574.1
39	*Prunus tenella*	Russian almond	Rosaceae	AEB69782.1
40	*Ziziphus jujuba*	Jujube	Rhamnaceae	XP_015901870.1
41	*Quercus suber*	Oak tree	Fagaceae	XP_023891811.1
42	*Durio ibethinus*	Coffee	Malvaceae	XP_022738256.1

RESULTS AND DISCUSSION

In order to study the nucleotide sequences, first the correct reading frame was obtained based on the CBF gene sequence in hexaploid wheat. After RT-PCR and making cDNA, PCR reaction was performed to amplify the CBF gene with specific primers. The result of PCR reaction on gel electrophoresis was examined and a band slightly larger than 600 bp, i.e. about 645 nucleotides, the size of CBF gene was determined (Figure 1). Examination of CBF transcription factor properties in hexaploid wheat plant shows that this gene is located on chromosome 5B (Park et al., 2018). The nucleotide length of this gene among the forty plants studied above ranged from 453 nucleotides for pine to 759 nucleotides for rice.

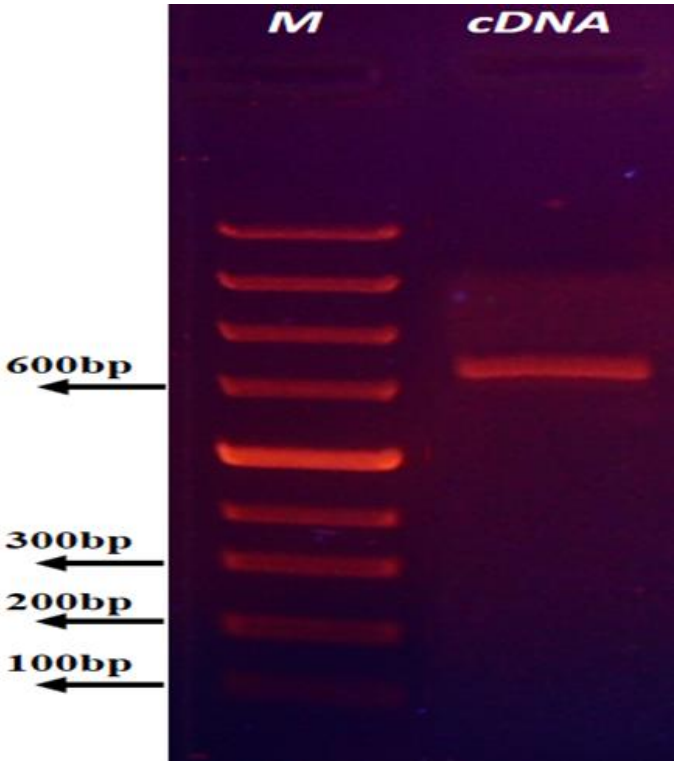

Figure 1. Image of electrophoresis gel and cDNA band, using CBF gene primers in wheat species (*Triticum turgidum*)

Protein studies of this gene, which were performed by bioinformatics analysis of its constituent amino acids, showed that the length of this gene in hexaploid wheat plant has 645 pairs of nucleotides and 214 amino acids. The isoelectric point and molecular weight of this protein were calculated to be 6.53 and 23597.80 kDa, respectively, which is one of the important properties of amino acids and based on which the polarity of the protein is determined. The atomic formula of this protein is C1043H1609N301O303S12. The most amino acids of this gene in tetraploid wheat plant consist of 13.6% alanine, 9.3% arginine and 8.4% leucine (Table 2). CBF protein is an intracellular protein that is also phosphorylable, and in an environment of high acidity due to the relatively high activity of proteolytic enzymes, the link between aromatic and hydrophobic amino acids is easily broken. This protein can also be glycosylated in the extracellular environment. Determination of amino acid levels is important in determining protein structure and cell evolution.

Table 2. The amino acids component of CBF protein

Amino Acid	Composition	%
Ala (A) 29	29	13.6
Arg (R) 20	20	9.3
Asn (N) 2	2	0.9
Asp (D) 13	13	6.1
Cys (C) 4	4	1.9
Gln (Q) 2	2	0.9
Glu (E) 14	14	6.5
Gly (G) 16	16	7.5
His (H) 5	5	2.3
Ile (I) 2	2	0.9
Leu (L) 18	18	8.4
Lys (K) 6	6	2.8
Met (M) 8	8	3.7
Phe (F) 10	10	4.7
Pro (P) 16	16	7.5
Ser (S) 16	16	7.5
Thr (T) 11	11	5.1
Trp (W) 7	7	3.3
Tyr (Y) 3	3	1.4
Val (V) 12	12	5.6
Sec (U) 0	0	0.0
Pyl (O) 0	0	0.0

Sequencing of this gene showed that there are differences in the sequences of bases of this gene in the regions of transcription initiation. The ABRE region with the CACGTG sequence, which is activated in response to abscisic acid, was found in the initial CBF gene sequence of most plants (Prerostova et al., 2020). The presence of the CAAT box, which is a specific region for increasing gene expression, was commonly seen in all plants. Alignment results of protein sequences for CBF gene showed that there is homology between the studied plants and most of them belong to the AP2 protected area. The strong presence of this highly conserved region with 60 to 70 amino acids was determined in the gene sequence of all studied plants. Which usually starts from nucleotide number 38 and ends at nucleotide number 97? However, in some samples, there were differences in the sequence of this region, which indicates the potential of plants containing this gene to cope with abiotic stresses and ultimately increase or decrease the expression of protein translated from the above gene. Careful study of these similar regions in the gene sequence in different plants allows the study of gene function, gene

expression and correlation with other genes in a particular pathway. Protected regions and this similarity in sequence can be studied to study the evolutionary relationships of CBF gene and as a special place in protein function and structure, as well as for primer design and identification of this gene in other nucleotide sequence plants in this region.

Bioinformatics analysis of amino acid sequences from CBF gene sequence translation showed that this gene is a member of the DREB gene subfamily due to its AP2 region (Yu et al., 2020). If the protein structure of CBF gene in hexaploid wheat indicates the presence of three major components. The AP2 region with three B-sheets and one α-helix chain. The pre-AP2 region contains the α-helix chain and Cystine transplantation. The section is after the AP2 area (Figure 2). The functional properties of the above proteins are largely derived from their three-dimensional structure, which is based on the homologous modeling of this protein (Figure 3). The spatial shape of the different parts of each protein actually determines the specific use and function of that part of the protein, and any change in this structure can disrupt its function (Mao and Hua, 2012). In this study, the three-dimensional structure of a part of the protein C-repeat binding factors was simulated. The molecular structure model of the protein shows that the region between amino acids 1 to 38 and 168 to 204 is one of the random regions of the protein coil, which includes the most irregularities. One of the reasons for the regular structure of proteins is the presence of the amino acids tyrosine, tryptophan, phenylalanine, valine and cysteine, which have the lowest levels in the structure of this protein (Table 2 and Figure 4). Other factors such as hydrophilicity, flexibility, molecular weight and isoelectric point can also be a reason for the irregular structure of this part of the protein (Holthauzen et al., 2011). It is likely that the greatest difference of this gene in the studied plants is related to the amino acid sequence of the above region. The presence of the majority of polar or hydrophobic amino acids in random coil regions indicates protein diversity at the cell surface.

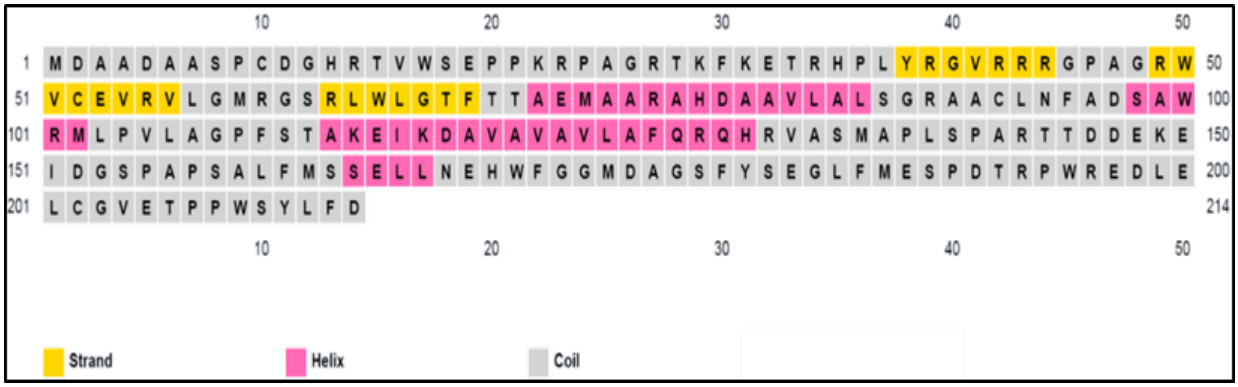

Figure 2. Secondary structure of C-repeat binding factors protein in terms of the presence of alpha helix, beta plates and random coil random

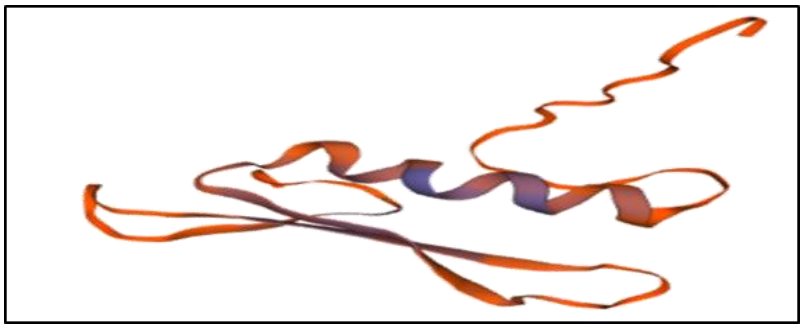

Figure 3. Part of the three-dimensional structure of C-repeat binding factors (Amino acid 88 to 214)

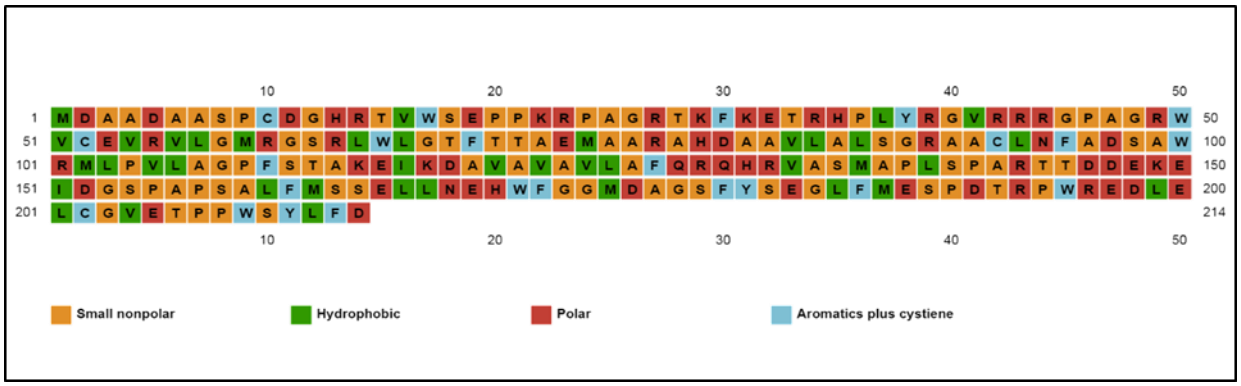

Figure 4. Secondary structure of C-repeat binding factors in terms of the presence of amino acids in polar, non-polar, hydrophobic and aromatic acids

In this study, the evolutionary matrix was calculated based on the Poisson correlation and the results of this matrix showed that the longest distance (4.324) there was between the protein sequence of CBF gene in sugar cane and nightshade and the lowest distance (0.007) was between bread hexaploid wheat and diploid wheat . Cluster analysis based on protein sequence of CBF gene in different plants showed that the studied plants are in three groups and subgroups (Figure 6 and 7).

According to the analysis of the cluster, it is possible to study the evolutionary relationships. In group one, the plants of the Gramineae family are included, although the plants of pine, grape and soybean were also included in this category separately. The largest distance in this group is between pine and rye (1.151) and rice (1.128). And if the lowest distance between plants of this group was in the Gramineae family and between diploid, tetraploid and hexaploid wheat (0.014). The results of this first class analysis confirm the existence of a relationship between the genus Triticum (wheat) and Aegilops (wheat). In the second group, more plants were placed than the other two groups (22 plants), so that a wide variety of distances were observed between the plants (Figure 6 and 7). In the meantime, there is the highest distance in terms of similarity between the protein sequence of saffron and ryegrass (1.106) and also the least amount of similarity between the sequence of the plant and Arabidopsis (0.025). Finally, only six plants were included in the third cluster, and all six plants belonged to completely separate families. The highest similarity distance is observed between strawberry and oil palm (0.589) and the lowest similarity distance is established between coconut and Asian palm (0.035). In comparison between the mentioned genera, sometimes a plant such as saffron with an extra-genus genus (tree species) is more similar in terms of this gene based on the Poisson model, which is one of the rare findings of this study. In order to better investigate the phylogenetic relationships based on CBF gene between the mentioned genera, the phylogenetic tree was drawn based on the method of maximum saving and with the help of plants in the NCBI database based on the sequence of this gene, the results of which can be seen in Figure 5. This comparison also showed that the sequence of CBF gene in Arabidopsis species is similar to what sequences of this gene are in the target species and other species.

Examination of the sequence of other stress-resistant genes in plants showed that the plants of the Gramineae family are very similar to each other and the analysis of kinship indicates the initial relationship and similarity of the gene sequence in plants of a family (Fatemi et al., 2015). Bioinformatics analysis of the nucleotide sequence of the CBF gene shows that the polymorphisms identified to differentiate the sequence of this gene, due to the addition or deletion of a certain number of nucleotides to the main sequence of this gene, are subject to natural selection. And the above polymorphisms can be considered as suitable molecular markers for the identification of CBF gene (Golovnina et al., 2007; Liu et al., 2019). On the other

hand, differences in nucleotide sequences and these polymorphisms can be due to different alleles of this gene. Genetic distance between two organisms is a justifiable difference between the two organisms using allelic diversity. In other words, genetic distance indicates the extent of genetic differences between populations or species that can be measured using some numerical quantities. Considering the similarity values of this gene among cultivars, it can be concluded that the cross between the cultivars that have the least similarity will have the greatest distance and the best result in achieving hybrids or achieving maximum differentiation in the generations after F1 (Noren et al., 2016; Ramezani and Rahimi, 2017). The higher alignment of the protein sequence of CBF gene among cold-resistant plants such as Gramineae family with protein sequence of AP2 region, which is the main factor of this gene, also indicates the high compatibility of these plants in cold resistance. The study of plant genetic diversity is a prerequisite for all plant breeding or conservation programs that show the relationship between molecular diversity and geographical diversity. The results of the phylogenetic tree indicate that there is usually less genetic distance between plants belonging to the same family in the study of this gene (CBF). The plants of the order, which were close to each other in terms of CBF gene sequence, are the order Gramineaes (wheat), the order Palmea (date palm) and the order Solanales (eggplant).

In fact, by examining the phylogenetic tree, it is possible to discover gene function, trace the origin of the gene, and identify the relatives of an organism (Naghavi et al., 2009).

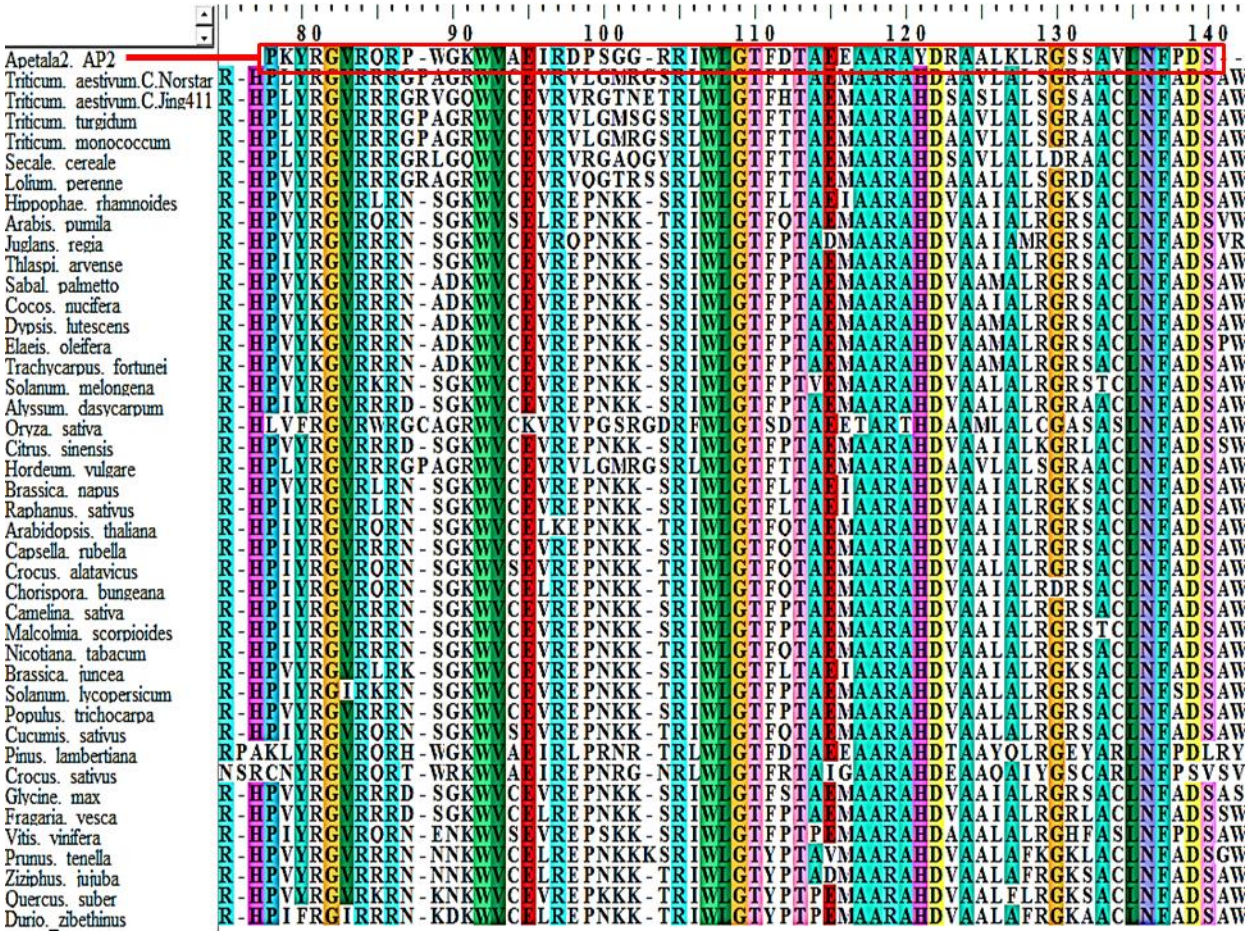

Figure 5. Aligning the protein sequence of the CBF gene in the studied plants, and aligning part of it with the AP2 region

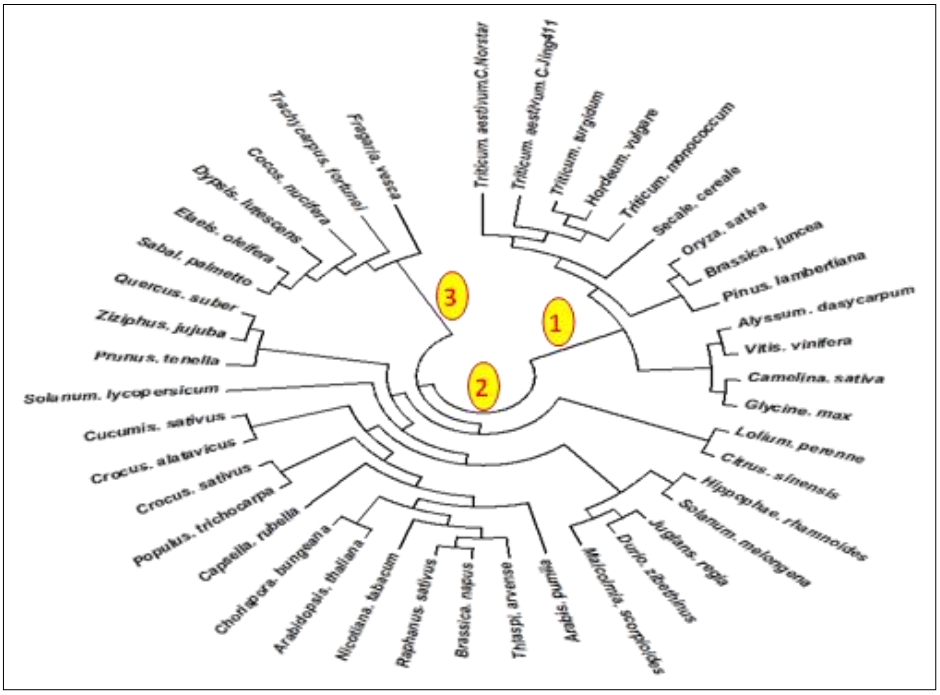

Figure 6. Cluster and cluster analysis of plants based on the variety of nucleotide sequences of the CBF gene

```
     1      2      3      4      5      6      7      8      9     10     11     12     13     14     15     16     17     18     19     20     21     22     23     24     25
26     27     28     29     30     31     32     33     34     35     36     37     38     39     40     41     42
 [ 1]
 [ 2] 2.1269
 [ 3] 0.0133 2.1269
 [ 4] 0.0066 2.1269 0.0066
 [ 5] 0.4426 2.4523 0.4529 0.4529
 [ 6] 2.1841 2.1269 2.1841 2.1841 2.3092
 [ 7] 2.6194 2.7147 2.6194 2.6194 2.3092 2.4523
 [ 8] 2.1841 2.6194 2.1841 2.1841 2.3782 2.3092 1.9262
 [ 9] 2.6194 2.3092 2.5324 2.6194 2.3782 2.7147 2.4523 2.2447
 [10] 2.3092 2.5324 2.3092 2.3092 2.3782 2.3092 1.6851 0.2382 2.2447
 [11] 2.5324 2.8201 2.5324 2.5324 2.3092 2.1841 2.3782 2.1841 2.1269 2.2447
 [12] 2.6194 2.6194 2.6194 2.6194 2.1841 2.1841 2.3092 2.3092 2.0215 2.2447 0.1345
 [13] 2.5324 2.8201 2.5324 2.5324 2.3092 2.1269 2.2447 2.2447 2.1269 2.3092 0.1046 0.1194
 [14] 2.5324 2.7147 2.5324 2.5324 2.3782 2.2447 2.3092 2.2447 2.1269 2.3092 0.0900 0.1421 0.1194
 [15] 2.6194 2.6194 2.6194 2.6194 2.1841 2.2447 2.3782 2.3782 1.9728 2.3092 0.1269 0.0405 0.1345 0.1575
 [16] 2.7147 2.6194 2.7147 2.7147 2.3092 2.3092 0.5064 2.3782 2.5324 2.3092 2.2447 2.2447 2.2447 2.1841 2.5324
 [17] 2.7147 2.1841 2.7147 2.7147 2.3782 2.5324 2.9378 2.8201 2.9378 2.6194 3.2255 3.0714 3.2255 3.2255 3.0714 2.8201
 [18] 1.9728 2.2447 2.0215 2.0215 2.2447 2.7147 2.3782 2.1269 2.6194 2.4523 2.3782 2.4523 2.3782 2.4523 2.5324 2.2447 3.2255
 [19] 2.3092 2.4523 2.3782 2.3782 2.2447 2.0215 2.1841 2.3092 2.1841 2.3092 2.3092 2.1269 2.3092 2.1841 2.2447 2.2447 3.0714 2.4523
 [20] 0.0544 2.0215 0.0544 0.0475 0.4740 2.1841 2.6194 2.3782 2.5324 2.3092 2.4523 2.6194 2.6194 2.5324 2.5324 2.8201 2.8201 2.1269 2.3782
 [21] 2.6194 2.6194 2.6194 2.6194 2.6194 2.4523 1.4909 0.3259 2.1841 0.1970 2.3782 2.3782 2.3782 2.3782 2.4523 2.3782 2.9378 2.4523 2.3092 2.6194
 [22] 2.5324 2.7147 2.5324 2.5324 2.6194 2.3782 1.4909 0.3259 2.1841 0.1497 2.3092 2.2447 2.3782 2.3092 2.3092 2.3782 2.8201 2.5324 2.3092 2.5324 0.0405
 [23] 2.6194 2.5324 2.6194 2.6194 2.4523 2.3782 1.7984 0.1890 2.3092 0.1970 2.2447 2.2447 2.3092 2.3782 2.3092 2.3782 2.7147 2.2447 2.3782 2.6194 0.2811 0.2637
 [24] 2.4523 2.6194 2.4523 2.4523 2.4523 2.1841 1.7214 0.1575 2.0728 0.1575 2.2447 2.2447 2.1269 2.3092 2.3092 2.3092 2.8201 2.3782 2.4523 2.5324 0.2466 0.2298 0.1421
 [25] 2.1269 2.3092 2.1269 2.1269 1.9262 2.5324 1.8818 1.6161 2.6194 1.6500 2.2447 2.3092 2.2447 2.2447 2.3782 2.5324 2.5324 2.3092 2.3092 2.1269 1.8392 1.8392 1.6500 1.6500
 [26] 2.5324 2.4523 2.5324 2.5324 2.4523 2.3782 1.7984 0.1810 2.3092 0.1890 2.2447 2.2447 2.3092 2.3782 2.3092 2.3782 2.7147 2.1841 2.4523 2.5324 0.2811 0.2637 0.0201 0.1345 1.6500
 [27] 2.1841 2.9378 2.2447 2.1841 2.3092 2.2447 2.7147 2.3092 2.5324 2.0728 2.4523 2.6194 2.3782 2.3782 2.7147 2.4523 2.3782 2.8201 2.2447 2.1841 2.1269 2.1269 2.1841 2.1841 2.6194 2.1841
 [28] 2.4523 2.4523 2.3782 2.4523 2.1841 2.4523 1.8392 1.5515 0.4847 1.6161 2.0215 1.9262 1.9262 2.0728 1.9728 2.5324 2.7147 2.3782 2.3782 2.6194 1.7592 1.7214 1.4338 1.3284 1.9262 1.4338
2.4523
 [29] 2.3092 2.4523 2.3092 2.3092 2.3782 2.1841 1.7214 0.1890 2.2447 0.1194 2.0728 2.1269 2.1841 2.2447 2.1841 2.1841 2.6194 2.1841 2.2447 2.3092 0.2551 0.2133 0.1345 0.1269 1.6161 0.1269
2.1269 1.5208
 [30] 2.1841 2.3782 2.1841 2.1841 2.2447 2.5324 2.4523 2.5324 2.7147 2.8201 2.1841 2.2447 2.2447 2.3092 2.3092 2.5324 3.2255 1.4338 2.3782 2.1841 2.8201 2.7147 2.6194 2.7147 2.3782 2.5324
2.4523 2.3092 2.6194
 [31] 2.7147 2.5324 2.7147 2.7147 2.9378 2.3782 2.1841 2.3782 2.1841 2.4523 2.4523 2.4523 2.3782 2.6194 2.3092 2.2447 2.6194 2.3092 2.6194 2.7147 2.5324 2.5324 2.4523 2.2447 2.3092 2.3782
2.1269 2.0215 2.1841 2.3782
 [32] 3.0714 2.6194 3.0714 3.0714 2.7147 2.4523 2.4523 2.2447 2.7147 2.2447 2.6194 2.8201 2.8201 2.7147 2.8201 2.8201 2.6194 2.3092 2.5324 3.0714 2.3782 2.3092 2.2447 2.3782 2.6194 2.2447
2.4523 2.4523 2.2447 2.7147 2.8201
 [33] 2.0728 2.3092 2.0728 2.0728 1.8818 2.6194 1.8392 1.6161 2.6194 1.6500 2.1841 2.2447 2.1841 2.1841 2.3092 2.4523 2.5324 2.3092 2.3092 2.0728 1.8392 1.8392 1.6500 1.6500 0.0133 1.6500
2.6194 1.9262 1.6161 2.3782 2.3092 2.6194
 [34] 2.2447 2.5324 2.3092 2.2447 2.3092 2.9378 2.7147 3.4078 2.9378 3.0714 3.2255 3.4078 3.4078 3.4078 3.4078 2.9378 2.5324 2.5324 2.8201 2.2447 3.0714 3.0714 3.6310 3.4078 3.4078 3.6310
2.7147 4.3241 3.4078 2.7147 2.7147 3.0714 3.4078
 [35] 3.0714 2.8201 3.0714 3.0714 2.5324 2.4523 2.6194 2.4523 2.7147 2.3782 3.2255 3.4078 3.2255 3.2255 3.4078 2.6194 2.6194 2.8201 2.9378 3.0714 2.2447 2.2447 2.3092 2.3782 2.6194 2.3782
2.7147 2.9378 2.3092 2.9378 2.5324 2.1841 2.6194 2.8201
 [36] 2.1269 2.7147 2.1841 2.1269 2.3092 2.2447 2.5324 2.0215 2.3782 2.0215 2.5324 2.6194 2.5324 2.6194 2.7147 2.7147 2.6194 2.3092 1.9262 2.0728 1.9262 1.9728 1.8818 1.9262 2.4523 1.8818
0.4634 2.1841 1.9262 2.7147 2.1841 2.4523 2.4523 3.0714 2.4523
 [37] 2.1841 2.3782 2.1841 2.1841 2.3092 2.3782 2.2447 2.2447 2.1841 2.6194 0.5286 0.5286 0.5286 0.5175 0.5514 2.5324 3.2255 2.5324 2.3092 2.1841 2.3782 2.3782 2.4523 2.4523 2.2447 2.4523
2.3782 2.1269 2.3782 2.0728 2.5324 2.5324 2.1841 2.9378 3.0714 2.2447
 [38] 1.9728 2.4523 2.0215 1.9728 2.1841 2.1841 2.4523 2.3092 2.1841 2.2447 2.3092 2.3092 2.2447 2.3782 2.2447 2.3782 2.6194 2.6194 2.0215 1.9262 2.1269 2.1841 2.3782 2.1841 2.4523 2.3782
0.6478 2.1841 2.3782 2.6194 2.1841 2.3782 2.4523 2.6194 2.6194 0.6865 2.7147
 [39] 2.3092 2.4523 2.3092 2.3092 2.3092 2.3092 2.2447 2.4523 2.3092 2.6194 2.1269 2.1269 2.0728 2.0215 2.1269 2.1269 3.0714 2.4523 2.2447 2.3092 2.5324 2.5324 2.4523 2.5324 3.0714 2.5324
2.1841 2.6194 2.3782 2.7147 1.9262 2.3782 3.0714 2.7147 2.3092 2.1841 2.3782 2.1269
 [40] 2.7147 2.9378 2.6194 2.6194 2.6194 3.0714 2.5324 2.7147 2.4523 2.6194 2.3782 2.4523 2.5324 2.3782 2.5324 2.4523 2.8201 2.3092 2.1841 2.6194 2.6194 2.7147 2.7147 2.7147 2.7147 2.7147
2.7147 2.6194 2.7147 2.7147 2.4523 2.8201 2.6194 3.4078 2.9378 2.2447 2.6194 2.8201 2.0728
 [41] 2.9378 3.0714 2.9378 2.9378 2.9378 2.7147 2.4523 2.4523 2.3092 2.6194 2.3092 2.2447 2.1841 2.2447 2.1841 2.2447 3.4078 2.4523 2.1841 3.0714 2.3782 2.5324 2.4523 2.4523 2.2447 2.5324
2.2447 2.2447 2.6194 3.6310 2.3782 2.9378 2.2447 2.9378 2.4523 2.1841 2.5324 2.4523 2.4523 2.0728
 [42] 2.8201 2.5324 2.8201 2.8201 2.3092 2.7147 2.5324 2.4523 0.5175 2.4523 1.9728 1.8818 1.9262 2.0728 1.8392 2.3782 2.4523 3.2255 2.3782 2.7147 2.5324 2.5324 2.5324 2.2447 3.2255 2.5324
2.3782 0.6106 2.4523 2.3092 2.1269 2.3782 3.2255 2.6194 2.4523 2.6194 2.2447 2.1841 2.8201 2.5324 2.9378
```

Figure 7. Cluster and cluster analysis of plants based on the variety of nucleotide sequences of the CBF gene

CONCLUSIONS

In general, the subfamilies Gramineae, Arecaceae and Solanaceae are classified into different groups and categories. Based on the obtained sequences, the study showed well that the method of searching for similarities and examining the protein expressed from the CBF gene, can correctly identify the sequence of this gene in genes for which no similar sequence has been available so far. Variety in the above sequences and in protein form is the basis of selections, and genotypic selection is also the basis of diversity. The existence of genetic diversity in this gene confirms that the differences in gene expression and differences in protein form are due to environmental effects, which in the long run have epigenetic effects and genetic modification of that gene in respective plant. The presence of the CBF gene, which is an important gene in cold stress, in different plants indicates the difference in the level of cold resistance of that plant, and the reason for this difference is the increase or decrease in the expression of this gene. If this distinction causes cold-sensitive cultivars. However, it is necessary to study other sensitive genes and possible linkage with the CBF gene.

ACKNOWLEDGEMENTS

The Department of Plant Production and the Research Department of Torbat Heydariyeh University are thanked for the scientific and financial support of this project.

REFERENCES

Baker, S.S., Wilhelm, K.S., Thomashow, M.F., 1994. *The 5' region of Arabidopsis thaliana COR15a has cis-acting elements that confer cold drought-and ABA-regulated gene expression.* Plant. Mol. Biol., 24: 701-713.

Eltayeb, A.E., Kawano, N., Badawi, G.H., Kaminaka, H., Sanekata, T., Shibahara, T., Inanaga, S., Tanaka, K., 2007. *Overexpression of monodehydroascorbate reductase in transgenic tobacco confers enhanced tolerance to ozone, salt and polyethylene glycol stresses.* Planta, 225: 1255-1264.

Esfandiari, E., Shekari, F., Esfandiari, M., 2007. *The effect of salt stress on antioxidant enzymes activity and lipid peroxidation on the wheat seedling.* Journal of Notulae Botanicae Horti Agribotanici Cluj-Napoca, 35: 48-56.

Fatemi, F., Najafi Zarrini, H., Heydari, P., 2015. *Phylogenetic and functional study of SOS1 gene to different plant species.* New Cellular and Molecular Biotechnology Journal, 5(17): 8-12.

Filippi, D.L., Fournier, M., Cameroni, E., Linder, P., Virgilio, C.D., Foti, M., Deloche, O., 2007. *Membrane stress is coupled to a rapid translational control of gene expression in chlorpromazine-treated cells.* Current Genetics, 52: 171-185.

Gilmour, S.J., Sebolt, A.M., Salazar, M.P., Everard, J.D., Thomashow, M.F., 2000. *Overexpression of the Arabidopsis CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation.* Plant Physiol., 124: 1854-1865.

Golovnina, K.A., Glushkov, S.A., Blinov, A.G., Adkison, L.R., Goncharov, N.P., 2007. *Molecular phylogeny of the genus Triticum.* Plant Syst. Evol., 264: 195-216.

Haake, V., Cook, D., Riechmann, J.L., Pineda, O., Thomashow, M.F., Zhang, J.Z., 2002. *Transcription factor CBF4 is a regulator of drought adaptation in Arabidopsis.* Plant Physiol., 130: 639-648.

Holthauzen, L.M., Auton, M., Sinev, M., Rösgen, J., 2011. *Protein stability in the presence of cosolutes.* Methods Enzymol., 492: 61-125.

Hsieh, T.H., Lee, J.T., Charng, Y.Y., Chan, M.T., 2002. *Tomato plants ectopically expressing arabidopsis CBF1 show enhanced resistance to water deficit stress.* Plant Physiology Research Article, 130: 618-626.

Jaglo, K.R., Kleff, S., Amundsen, K.L., Zhang, X., Kaake, V., Xhan, J.Z., 2001. *Components of the Arabidopsis C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in Brassica napus and other plant species.* Plant Physiology Research Article, 127: 910-917.

Jaglo-Ottosen, K.R., Gilmour, S.J., Zarka, D.G., Schabenberger, O., Thomashow, M.F., 1998. *Arabidopsis CBF1 overexpression induces COR genes and enhances freezing tolerance.* Science, 280: 104-106.

Liu, Y., Dang, P., Liu, L., He, C., 2019. *Cold acclimation by the CBF-COR pathway in a changing climate: Lessons from Arabidopsis thaliana.* Plant Cell Reports, 38: 511-519.

Loggini, B., Scartazza, A., Brugnoli, E., Navari-Izzo, F., 1999. *Antioxidative defense system pigment composition and photosynthetic efficiency in two wheat cultivars subjected to Drought.* Plant Physiology, 119: 1091-1100.

Mao, X., and Hua, Y., 2012. *Composition, structure and functional properties of protein concentrates and isolates produced from walnut (Juglans regia L.).*

International Journal of Molecular Sciences, 13: 1561-1581.

Maurya, J.P., and Bhalerao, R.P., 2017. *Photoperiod- and temperature-mediated control of growth cessation and dormancy in trees: a molecular perspective.* Annals of Botany, 120: 351-360.

Medina, J., Bargues, M., Terol, J., Perez-Alonso, M., Salinas, J., 1999. *The Arabidopsis CBF gene family is composed of three genes encoding AP2 domain-containing proteins whose expression is regulated by low temperature but not by abscisic acid or dehydration.* Plant Physiology, 119: 463-470.

Naghavi, M.R., Malboobi, M.A., Rashidi, S., 2009. *Bioinformatics.* University of Tehran Press, Iran. (In Persian)

Nanjo, T., Kobayashi, M., Yoshiba, Y., Kakubari, Y., Yamaguchi-Shinozaki, K., Shinozaki, K., 1999. *Antisense suppression of proline degradation improves tolerance to freezing and salinity in Arabidopsis thaliana.* FEBS Letters, 461: 205-210.

Noren, L., Kindgren, P., Stachula, P., Ruhl, M., Eriksson, M.E., Hurry, V., 2016. *Circadian and plastid signaling pathways are integrated to ensure correct expression of the CBF and COR genes during photoperiodic growth.* Plant Physiology, 171: 1392-1406.

Novillo, F., Alonso, J.M., Ecker, J.R., Salinas, J., 2004. *CBF2/DREB1C is a negative regulator of CBF1/DREB1B and CBF3/DREB1A expression and plays a central role in stress tolerance in Arabidopsis.* Proceedings of the National Academy of Sciences USA, 87: 291-293.

Park, S., Lee, C.M., Doherty, C.J., Doherty, C.J., Gilmour, S.J., Kim, Y., Thomashow, M.F., 2015. *Regulation of the Arabidopsis CBF regulon by a complex low-temperature regulatory network.* Plant Journal, 82: 193-207.

Park, S., Gilmour, S.J., Grumet, R., Thomashow, M.F., 2018. *CBF-dependent and CBF-independent regulatory pathways contribute to the differences in freezing tolerance and cold-regulated gene expression of two Arabidopsis ecotypes locally adapted to sites in Sweden and Italy.* Plose one, 10: 1371.

Prerostova, S., Černý, M., Dobrev, P., Motyka, V., Hluskova, L., Zupkova, B., Gaudinova, A., Knirsch, V., Janda, T., Brzobohatý, B., Vankova, R., 2020. *Light regulates the Cytokinin-Dependent cold stress responses in Arabidopsis.* Frontiers in Plant Science, 11: 336-350.

Ramezani, M., and Rahimi, M., 2017. *Grouping and estimation of genetic diversity of different ecotypes of medicinal plant of Plantago psyllium using ISSR marker.* Journal of Molecular and Cellular Research, 30(2): 312-322.

Shi, Y., Huang, J., Sun, T., Wang, X., Zhu, C., Ai, Y., 2017. *The precise regulation of different COR genes by individual CBF transcription factors in Arabidopsis thaliana.* Journal Integrative Plant Biology, 59: 118-133.

Shi, Y., Ding, Y., Yang, S., 2018. *Molecular regulation of CBF signaling in cold acclimation.* Trends Plant Sciences, 23: 623-637.

Stockinger, E.J., Gilmour, S.J., 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.* Proceedings of the National Academy of Sciences USA, 94: 1035-1040.

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., Yamaguchi-Shinozaki, K., Shinozaki, K., 2002. *Important roles of drought-and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana.* Plant Journal, 29: 417-426.

Thomashow, M.F., 2001. *So what's new in the field of plant cold acclimation.* Lots Plant Physiol., 125: 89-93.

Vaidyanathan, H., Sivakumar, P., Chakrabarsty, R., Thomas, G., 2003. *Scavenging of reactive oxygen species in NaCl-stressed rice (Oryza sativa L.) differential response in salt-tolerant and sensitive varieties.* Plant Science, 165: 1411-1418.

Xin, Z., and Browse, J., 2000. *Cold comfort farm: the acclimation of plants to freezing temperatures.* Plant Cell Environ., 23: 893-902.

Xin, Z., Mandaokar, A., Chen, J., Last, R.L., Browse, J., 2007. *Arabidopsis ESK1 encodes a novel regulator of freezing tolerance.* The Plant Journal, 49: 786-799.

Zhao, C., Zhang, Z., Xie, S., Si, T., Li, Y., Zhu, J.K., 2016. *Mutational evidence for the critical role of CBF transcription factors in cold acclimation in Arabidopsis.* Plant Physiology, 171: 2744-2759.

Zhen, Y., and Ungerer, M.C., 2008. *Relaxed selection on the CBF/DREB1 regulatory genes and reduced freezing tolerance in the southern range of Arabidopsis thaliana.* The Society for Molecular Biology and Evolution, 25: 2547-2555.

Yu, H., Kong, X., Huang, H., Wu, W., Park, J., Yun, J., Lee, B., Shi, H., Zhu J., 2020. *STCH4/REIL2 confers cold stress tolerance in Arabidopsis by promoting rRNA processing and CBF protein translation.* Cell Reports, 30(1): 229-242.