UNIVERSITÉ DU QUÉBEC À MONTRÉAL

FONCTIONS GÉNOMIQUES DE LA FAMILLE DES FACTEURS DE TRANSCRIPTION AP2 CHEZ LES CÉRÉALES

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN BIOLOGIE

PAR

MOHAMED BADAWI

AVRIL 2008

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

FUNCTIONAL GENOMICS OF THE AP2 TRANSCRIPTION FACTOR FAMILY IN CEREALS

PH.D. THESIS

PRESENTED

AS A PARTIAL REQUIREMENT

FOR THE DOCTORATE IN BIOLOGY

PAR

MOHAMED BADAWI

APRIL 2008

To soul of my parents

REMERCIEMENTS

My sincere thanks to my supervisor Dr. Fathey Sarhan for giving me an opportunity to work in his laboratory. I am grateful for his guidance and stimulating discussions. I am particularly thankful for his constant support and vision throughout my tenure as a doctoral student in his lab. I will always be indebted to him for introducing me to molecular biology and lab techniques.

I would like to thank Prof. Jean Danyluk for his valuable advices and for taking special interest in my research work. He has provided me much scope for thought during our meetings and advisory sessions. Gratitude also goes to Dr. Zahra Agharbaoui for her various suggestions and help in operations and also for the encouragement during the thesis research.

I am very grateful to the ministry of state and higher education and scientific research in Egypt for their financial support during my PhD studies.

I also thank my wonderful friends and colleagues in the laboratory for their unwavering encouragement that has been the most important factor in the completion of this task.

Finally, I would like to thank my family back home, for their undaunting perseverance and support all this while that I have been away from them in the pursuit of my goal.

TABLE DES MATIÈRES

LISTE DES FIGURES

Article I

Article II

Article III

LISTE DES TABLEAUX

Article I

Article II

Article III

LISTE DES ABRÉVIATIONS, DES SIGLES ET DES ACRONYMES

ABA	:	abscisic acid
aa	:	amino acid
AP2	:	APETAL2
bHLH	:	basic helix loop helix
At	:	*Arabidopsis thaliana*
ATP	:	adenosine triphosphate
bp	:	base pairs
cDNA	:	complementary DNA
°C	:	degrees Celsius
CaMV	:	Cauliflower Mosaic Virus
COR	:	Cold regulated
CRT	:	C-repeat
C-terminus	:	Carboxy terminus
Da	:	Dalton
dCTP	:	Desoxycytidine triphosphate
DRE	:	Dehydration responsive element
DREB	:	DRE binding protein
DNA	:	deoxyribonucleic acid
E. coli	:	*Escherichia coli*
GST	:	Glutathione-S-Transferase
GFP	:	green fluorescent protein
hr	:	hour
kb	:	kilobase
LEA	:	late embryogenesis abundant
min	:	minute

mM	:	millimolar
mRNA	:	messenger RNA
MS	:	Murashige and Skoog
NCBI	:	National Center for Biotechnology Information
N-terminal	:	amine-terminus
PAGE	:	polyacrylamide gel electrophoresis
PCR	:	polymerase chain reaction
PVDF	:	polyvinylidene difluoride
QTL	:	quantitative trait loci
RNA	:	ribonucleic acid
RNase	:	ribonuclease
RT	:	reverse transcription
SDS	:	sodium dodecyl sulphate
TBE	:	Tris-borate-EDTA buffer
v/v	:	volume/volume
WT	:	wild type

RÉSUMÉ

Les plantes sont exposées, à chaque étape de leur cycle de vie, à différents stress abiotiques notamment la sécheresse, le froid et la salinité. Bien que ces différents types de stress aient des effets spécifiques pour chacun d'entre eux, ils ont tous la capacité d'induire un effet commun soit le déficit hydrique qui affecte la croissance et le développement des plantes. Certaines plantes s'adaptent au déficit hydrique en modifiant leurs métabolismes ainsi que l'expression de leurs gènes afin de synchroniser leur développement avec les conditions de stress. Cette modification se traduit par la synthèse de nouvelles protéines, particulièrement régulée par le déficit hydrique, notamment les protéines de type Ap2/EREBP. Chez *Arabidopsis thaliana*, ces protéines forment une famille de facteurs de transcription de 145 membres. Certains membres jouent des rôles essentiels de régulateur dans plusieurs processus de développement tandis que d'autres sont impliqués dans les mécanismes de réponses des plantes aux stress. Chez les céréales, très peu de gènes *Ap2/EREBP* avec des fonctions connues ont été identifiés. Ce projet de recherche vise à identifier cette famille de gènes chez le blé hexaploid (*Triticum aestivum* L.). Le choix de ce type de blé comme modèle d'étude réside dans sa capacité de tolérance au gel et sa grande variabilité de réponses aux stress abiotiques. Le blé est la céréale la plus importante au monde puisqu'il représente environ 73% de la production céréalière mondiale. Il est cultivé dans divers types d'environnements climatiques, de productions et de fermes. De plus, il représente la principale source d'énergie alimentaire, d'emploi et de revenu dans plusieurs pays en voie de développement. Le blé agronomique est plus nutritif et économique qu'*Arabidopsis*. Cependant, le choix entre les deux espèces est influencé par la taille et la complexité de leurs génomes.

Le premier chapitre décrit l'identification, la phylogénie et la localisation chromosomique des gènes *Ap2/EREBP* du blé. Une approche génomique combinant une recherche de bases de données de séquences suivie d'une analyse bioinformatique, un criblage de banques d'ADNc et des amplifications par réaction de polymérisation en chaîne a permis d'identifier 107 ADNc du blé qui codent potentiellement pour des facteurs de transcription de type Ap2/EREBP. Leur localisation chromosomique montre que ces gènes sont dispersés sur tout le génome du blé. Cependant, quelques gènes de cette famille, appelés CBFs (C-repeat Binding Factors), sont regroupés principalement sur le long bras du groupe de chromosomes 5 et ont été associés directement à des déterminants génétiques (QTLs) contrôlant les réponses aux stress et divers aspects du développement floral.

Le deuxième chapitre présente la caractérisation de la sous-famille de gènes CBFs. L'analyse phylogénétique a indiqué que les espèces de blé contiennent au moins 23 gènes différents de type CBF. Chez les *Poaceae*, les CBFs sont classés dans 10 groupes qui partagent une origine phylogénétique commune et des caractéristiques structurales semblables. Six de ces groupes (3C, 3D, 4A, 4B, 4C et 4D) sont trouvés

seulement chez les *Pooideae*, suggérant ainsi qu'ils représentent les mécanismes de réponse de CBF qui ont évolué récemment pendant la colonisation des habitats tempérés. Les études menées sur le profil d'expression de ces gènes démontrent que 5 des groupes spécifiques aux *Pooideae* montrent une expression constitutive élevée et une expression inductible par les basses températures chez le cultivar d'hiver. Par contre à des températures modérées, la régulation de l'expression varie en fonction de la période de la journée. L'expression inductible et non héritée au sein des groupes de CBF a possiblement joué un rôle prédominant dans l'aptitude des cultivars d'hiver à tolérer les basses températures et elle est probablement à la base de la variabilité génétique de la tolérance au gel chez les *Pooideae*.

Le troisième article décrit l'identification de deux gènes chez le blé, ICE1-like (Inducteur de l'expression de CBF 1) codant pour un facteur de transcription qui régule l'expression des gènes CBF. *TaICE87* et *TaICE41* codent pour un activateur de transcription de type MYC-like bHLH. *TaICE87* et *TaICE41 se lient spécifiquement à différentes* séquences de reconnaissance MYC du promoteur de *TaCBFIVd-B9*. *TaICE87* and *TaICE41* sont constitutivement exprimés chez le blé et leur surexpression chez *Arabidopsis* mène à l'augmentation du niveau d'expression des gènes *AtCOR* et *AtCBF3* et de la tolérance des plantes au gel. Ces résultats suggèrent que *TaICE87* and *TaICE41 sont des* orthologues fonctionnels d'AtICE1 et peuvent réguler la transcription d'AtCBF3. La structure complexe des éléments MYC au niveau des promoteurs CBF et la différence d'affinité entre TaICE87 et TaICE41 pour les éléments MYC suggèrent que ces 2 protéines peuvent différentiellement activer CBF chez le blé.

En résumé, le clonage de la famille de gènes *Ap2/EREBP* chez le blé a permis d'identifier et de caractériser les membres CBF régulés par les basses températures. De plus, les résultats de cette étude démontrent que l'expression des gènes CBF est amplifiée chez les *Pooideae*, suggérant ainsi un rôle plausible dans la tolérance au froid des cultivars d'hiver. L'analyse des gènes *CBFs* pourrait nous conduire à améliorer la tolérance au gel chez le blé, et par conséquent faciliter l'établissement de nouvelles stratégies visant à améliorer la productivité des céréales et leur adaptation aux changements climatiques.

Mots clés : blé, acclimatation au froid, tolérance au gel, facteurs de transcription, Ap2/EREBP

SUMMARY

Crop plants are exposed to many types of abiotic and biotic stresses during their life cycle. Water deficit caused by drought, low temperature or high salt concentration in the soil, is one of the most common environmental stress that affect growth and development of plants. These environmental factors limit the geographical distribution and growing season of many plant species, and affect crop quality and productivity.

A number of reports have shown that these stresses can alter plant metabolism and gene expression. The *AP2/EREBP* transcription factor gene family is plant specific regulatory genes modulated by different stresses. In *Arabidopsis* the AP2/EREBP comprises a large and diverse transcription factor family of 145 members. Members of this superfamily have been shown to play a variety of roles such as being key regulators of several developmental processes and being involved in plant response mechanisms to various types of biotic and environmental stresses.

In cereals, very few *AP2/EREBP* genes with known functions have been identified. In order to characterize the *AP2/EREBP* genes in cereals and to elucidate their roles and regulation during abiotic stress, the identification of this family of genes was undertaken in wheat. Hexaploid wheat was selected as a study model for freezing tolerance (FT) and other abiotic stresses. Wheat is cultivated in a wide range of climatic environments, production environments and farming systems. In addition, it constitutes a principal cereal crop that provides nutritional energy and is the first source of protein in developing countries. Agriculturally, wheat is more nutritive and economic than *Arabidopsis*. However, this choice is accompanied by greater challenges because of the size and the complexity of the wheat genome.

The results are presented and discussed in three articles. In the first article, the identification, phylogeny and chromosomal localization show that *AP2/EREBP* factors are coded and functional on wheat genome. Using a genomic approach in combination with bioinformatics analysis and molecular studies allowed us to identify 107 *AP2/EREBP* factors from hexaploid wheat (*Triticum aestivum* L.). The wheat *AP2/EREBP* cDNAs were identified through a search in different data bases, PCR or cDNA screening. However, some genes of this family called CBFs (C-repeat Binding Factors) are cluster mainly on the long arm of chromosomes 5 and directly connected to genetic QTLs (Quantitative Trait Loci) determinants controlling the responses to environmental and biotic stresses, in addition to various aspects of the floral development.

In the second article we initiated a study to identify and characterize CBF transcription factors from hexaploid wheat. The *Poaceae* is divided into several subfamilies. (The *Oryzaceae* (rice), *Panicoideae* (maize) have a more tropical geographical distribution compared to members of the Pooideae which contain the temperate cereals wheat, barley and oat.) Our analyses revealed that wheat species

contain at least 23 different CBF genes, and that the *Poaceae* CBFs are classified into 10 groups that share a common phylogenetic origin and similar structural characteristics. Six of these groups (3C, 3D, 4A, 4B, 4C and 4D) are found only in the *Pooideae*, suggesting that they represent the CBF response machinery that evolved recently during the colonization of temperate habitats. Expression studies revealed that 5 of the *Pooideae*-specific groups display higher constitutive and low temperature inducible expression in the winter cultivar, and a diurnal type regulation during growth at warm temperature. The higher inherited and inducible expression within these CBF groups may play a predominant role in the superior low temperature tolerance capacity of the winter cultivars and could possibly be at the basis of the genetic variability in freezing tolerance within the *Pooideae* subfamily.
The third article reports the identification of two *ICE1*-like genes from wheat (*inducer of CBF expression 1*), encoding an upstream transcription factor that regulates the transcription of *CBF* genes. TaICE87 and TaICE41 encode MYC-like bHLH transcriptional activator. *TaICE87* and *TaICE41* bind specifically to different MYC recognition sequences in *TaCBFIVd-B9* promoter. *TaICE87* and *TaICE41* are expressed constitutively, and their overexpression in *Arabidopsis* leads to increased expression of *AtCBF3* and *AtCOR* genes and increased tolerance to freezing stress. These results suggest that TaICE87 and TaICE41 are functional ortholog of AtICE1 and can regulate the transcription of AtCBF3. The complex structure of MYC elements in wheat CBF promoters and the different affinities of TaICE87 and TaICE41 for MYC elements suggest that these proteins might activate different CBF in wheat.
In conclusion the cloning of the *AP2/EREBP* genes family from wheat helps the identification and characterization of low-temperature (LT)-regulated members (CBF) in this family. This research showed that CBF amplified in *Pooideae* subfamily and might play a predominant role in the superior LT tolerance of winter cultivars. The function analysis of CBFs genes may be used as a tool to improve stress tolerance in wheat. The elucidation of the regulatory mechanisms that control FT trait in complex systems like wheat will facilitate the development of effective engineering strategies to increase stress tolerance and cereal productivity. This will help in breeding new cultivars that adapt to climate changes.

Key words: wheat, cold acclimation, freezing tolerance, transcription factor, *Ap2/EREBP*

INTRODUCTION AND LITERATURE REVIEW

Plants are exposed to various adverse environmental conditions such as drought, high salt and high/low temperature during their life cycle. These environmental factors severely limit plant growth and productivity. One of the plant responses to these stresses is the expression of a large number of genes, whose products are known or believed to be involved in various adaptive functions under stress conditions (Thomashow 1999). Genes induced during drought and cold stress conditions encode proteins that function not only in the protection of cells from stress, but also in the gene expression and signal transduction in stress response (Thomashow, 1999; Shinozaki and Yamaguchi-Shinozaki, 2000; Zhu, 2002; Shinozaki et al 2003). To understand the nature and the function of the genes involved in stress tolerance, an intensive research was dedicated to the identification and characterization of stress-inducible genes. The products encoded by the genes identified can be classified into two groups based on their presumed function (Breton et al 2000) (Fig. 1 Shinozaki and Yamaguchi-Shinozaki 2006). The first group includes proteins that probably function in stress tolerance, such as chaperones, late embryogenesis abundant (LEA) proteins, osmotin, antifreeze proteins, mRNA-binding proteins, key enzymes for osmolyte biosynthesis, water channel proteins, sugar and proline transporters, detoxification enzymes and various proteases. The second group contains protein factors involved in further regulation of signal transduction and gene expression that probably function in stress response, such as protein kinases, transcription factors (TFs) and enzymes involved in phospholipid metabolism. The existence of a variety of stress-inducible genes suggests complex responses of plants to abiotic stress.

1. Structural and functional proteins

1.1 Antifreeze proteins (AFPs)

AFPs were first identified in fishes as the causative agents of serum freezing point depression (DeVries 1971). These proteins are unique in that they cause a freezing point depression far greater than would be predicted from colligative properties alone. AFPs were shown to lower the freezing point of water by interacting directly with the ice surface and inhibiting the binding of additional water molecules to the ice crystal lattice (Breton et al 2000). AFPs are also known to inhibit the recrystallization of ice, which is the growth of larger ice crystals at the expense of smaller ice crystals. Larger ice crystals increase the possibility of physical damage within frozen tissues. The inhibition of the recrystallization of ice occurs at very low AFP concentration (nM) and may be the function of AFPs in freezing tolerant organisms. In addition, they may be involved in protecting cell membranes from cold induced damage (Breton et al 2000). Many types of AFPs have been identified in fish, insects and plants (Huang et al 2002a). However their composition and structure are varied. The levels of thermal hysteresis range from 0.2-1.6°C for plant AFPs, 1-2°C for fish AFPs to 5-10°C for insect AFPs. Transgenic *Arabidopsis* plants which express several genes encoding insect (*Dendroides canadensis*) antifreeze proteins did not demonstrate improved ability to survive freezing when compared to the wild type (Huang et al 2002a).

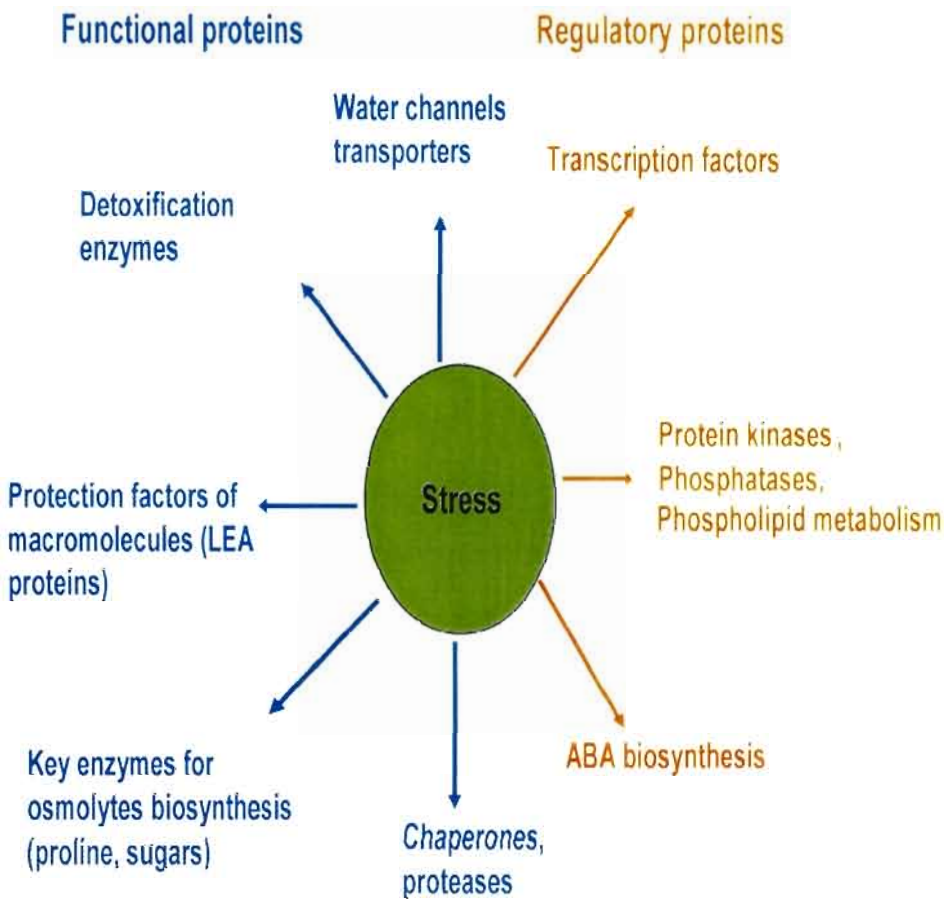


Figure1: Functions of drought and cold stress-inducible genes in stress tolerance and response. Gene products are classified into two groups. The first group includes proteins that probably function in stress tolerance (functional proteins), and the second group contains protein factors involved in further regulation of signal transduction and gene expression that probably function in stress response (regulatory proteins) (by Shinozaki and Yamaguchi-Shinozaki 2006).

1.2 COR/LEA proteins

It is well established that freeze-induced membrane damage results primarily from the severe dehydration associated with freezing (Thomashow 1999). Furthermore, the dehydration stress that occurs during freezing is likely to be similar to the one occurring during drought and salt stress, and so is expected to trigger some common genetic and physiological responses. LEA proteins were initially found to be induced during late embryogenesis, prior to seed desiccation. They were divided into different groups based on conserved structural features (Thomashow 1999). The hydrophilic COR proteins of unknown functions are referred to as COR/LEA proteins. The function of the LEA gene products is not clearly defined yet, but they are considered to have cellular protective functions. LEA proteins are highly hydrophilic, and repeating units composed of short tracts of conserved amino acids are commonly found in them (Thomashow 1999). Thus, it is speculated that these proteins retain water molecules and prevent crystallization of cellular components under water deficit, which results from drought, high salt and freezing stresses. Indeed, there is experimental evidence showing that overexpression of a LEA gene results in enhanced drought, salt and freezing tolerance (Breton et al 2000). Transgenic rice plants that overexpress a group III LEA gene from barley showed increased tolerance to water deficit and salinity as reflected by their higher growth rates, delayed development symptoms and improved recovery (Xu et al 1996). However, the effect of cold stress on the performance of these transgenic plants was not measured.

1.3 Genes involved in osmotic adjustment and detoxification

When plants are exposed to abiotic stresses such as LT, membrane fluidity and membrane-bound processes are perturbed (Bohnert, et al 1999). In addition, when extracellular ice forms, there is a flux of water from the cells to the apoplast causing

both a dehydration of the protoplast and an increase in the concentration of intracellular solutes, thus putting a strain on membranes and macromolecules (Delauney and Verma 1993). Under conditions in which chloroplasts are exposed to excess excitation energy generated by the interaction of light and LT, there is an increase in photoreduction of oxygen and concomitant production of reactive oxygen intermediates such as superoxides and peroxides, which can damage membranes and enzymes. Thus, many compounds like osmoprotectants serve to raise osmotic pressure in the cytoplasm and can also stabilize proteins and membranes when salt levels or temperatures are unfavorable. Osmoprotectants therefore play important roles in the adaptation of cells to various adverse environmental conditions (Yancey, 1994). These osmoprotectants, known as compatible osmolytes since they do not inhibit other cellular functions, include amino acids (proline, ectoine), quaternary amines (glycine betaine, alanine betaine) and sugars (mannitol, pinitol) among others. The enzymes that are involved in the synthesis of these compounds, like betaine aldehyde dehydrogenase, are induced (Weretilnyk and Hanson, 1989). Genes encoding enzymes that participate in the removal of toxic intermediates produced during stress responses such as glutathione-S-transferase, glutathione reductase and superoxide dismutase are also induced by stresses (Hayes and Pulford 1995). These enzymes participate in the removal of reactive oxygen species generated by abiotic stresses. In addition, genes encoding proteases, ubiquitin extension protein, chaperonins and small heat-shock proteins are also expressed in response to abiotic stresses (Vierstra 1996). These proteins are probably involved in the removal or repair of damaged cellular components.

2. Genes of regulatory and signaling molecules

Involvement of various kinases and phosphtases in stress signal transduction has been well documented by numerous genetic and biochemical studies (Shinozaki

and Yamaguchi-Shinozaki 1997). Many of the corresponding genes, including Ca^{2+} - dependent protein kinases and mitogen-activated protein kinases are known to be induced by drought, cold and high salt (Shinozaki and Yamaguchi-Shinozaki 1997). Genes encoding signaling enzymes such as G proteins and phospholipase C are also responsive to abiotic stresses (Shinozaki and Yamaguchi-Shinozaki 1997). In addition, many genes encoding transcription factors such as WRKY, MYB, MYC or proteins that possess the structures like homedomain, basic leucine zipper (bZIP), AP2 domain and zinc finger proteins are upregulated by various stresses. WRKY proteins bind to the W box sequence found in the promoter of the pathogen-responsive gene chitinase CHN50 in tobacco. WRKY proteins have been characterized in diverse plant species (*Arabidopsis*, parsley, and tobacco) (Chakravarthy et al 2003). The bZIP family of TFs includes TGA and GBF factors and has been characterized in tobacco, soybean, and *Arabidopsis*. They bind to the as-1 and G box *cis* elements, respectively. The as-1 element was shown to be responsive to the defense signaling molecules salicylic acid and jasmonate. The G box is a ubiquitous element, and it has been proposed that it functions in concert with neighboring *cis* elements in regulating gene expression related to different functions, including pathogen attack. The MYB family of TFs is a very large family with a subset of genes that play a role in defense response. MYB family members bind to several different *cis*-element sequences. Promoter analysis of structural genes induced by abiotic and biotic stresses revealed the presence of *cis*-elements conferring inducibility to stresses (Chakravarthy et al 2003). The *cis*-element called C-repeat or drought responsive element (DRE), was found to mediate dehydration, high salt and cold regulation of the gene during abiotic stress while the *cis*-element GCCGCC called the GCC box was found to mediate the regulation of PR genes during biotic stress. A DRE, TACCGACAT, has been identified in the rd29A promoter (Yamaguchi-Shinozaki and Shinozaki 1994). The core sequence CCGAC element is also present in many cold-regulated genes (Baker et al 1994). Transcription factors interacting with the C-repeat/DRE and the GCC box have been isolated from *Arabidopsis* (Stockinger et al 1997). Thomashow's group isolated a C-repeat/DRE

Binding Factor CBF1 using the yeast one-hybrid strategy (Stockinger et al 1997). The CBF factor contains an AP2 DNA-binding domain present in several plant species. Later, the factor and other CBF family members referred to as DREBs (DRE Binding Proteins) have also been reported by another group (Liu et al 1998). Overexpression of CBF1 resulted in the expression of several cold-inducible genes and enhanced freezing tolerance (Jaglo-Ottosen et al 1998). Similarly, overexpression of the DREB1A protein resulted in enhanced freezing and drought tolerance (Liu et al 1998). CBF/DREB proteins function in ABA-independent cold and drought/high salt signaling pathways. Haake et al (2002) showed that CBF4 plays a role during drought adaptation. In contrast to the three already identified CBF/DREB1 homologs, which are induced under cold stress, CBF4 gene expression is up-regulated by drought stress, but not by low temperature.

3. AP2/EREBP Transcription factors

Transcription factors are important regulators of gene expression. In general, a transcription factor is composed of at least two discrete domains, a DNA-binding domain and an activation/repression domain, which operate together to regulate many physiological and biochemical processes by modulating the rate of transcription initiation of target genes (Ptashne 1988). It is not surprising to note that the plants assign a great portion of their genome to the transcription factors: over 1600 genes (6% of the genome) were identified in *Arabidopsis thaliana* (Riechmann et al 2000) and in rice (Goff et al 2002). According to the type of DNA-binding domains, eukaryotic transcription factors are classified in several families. During the last decade, it was shown that the transcription factors of the AP2 family play a major role in the genetic control of floral development and the response to biotic and abiotic stresses.

3.1 Definition of AP2/EREBP

The AP2/EREBP (APETALA2)/EREB (Ethylene Responsive Element Binding Factor) domain is a DNA-binding domain that consists of approximately 60 conserved amino acid residues (Jofuku et al 1994 and Hao et al., 1998). The AP2 domain was first recognized as a repeated motif within the *Arabidopsis thaliana* AP2 protein (Jofuku et al 1994). Shortly afterwards, four DNA-binding proteins from tobacco were identified that interact with a sequence that is essential for the responsiveness of some promoters to the plant hormone ethylene, and were designated as ethylene-responsive element binding proteins (EREBPs) (Ohme-Takagi and Shinshi, 1995). The genes containing a single or two AP2/ EREBP domains encode putative transcription factors and belong to the AP2/EREBP multigene family (Riechmann and Meyerowitz 1998). Sakuma et al (2002) identified 145 AP2/EREBP genes in the genome of *Arabidopsis thaliana.*

3.2 Classification of AP2/EREBP gene family

The distinguishing characteristic of proteins of the AP2/EREBP family is that they contain either one or two *APETALA2 (AP2)* domains. The DNA-binding domain of the EREBP-2 was mapped to a region that was common to all four proteins (Ohme-Takagi and Shinshi, 1995), and was found to be closely related to the AP2 domain but that did not bear sequence similarity to previously known DNA-binding motifs. *Ap2/EREBP* genes form a large family, with many members known in several plant species. The *Arabidopsis* genome encodes 145 AP2-related proteins sharing a conserved DNA-binding domain of approximately 60 amino acids in length. Based on sequence similarities in their AP2 DNA binding domain, these proteins are classified into five groups: AP2 subfamily involved in development (14 genes, two AP2 domains), RAV subfamily (6 genes, one AP2 and one B3 domain), DREB

subfamily involved in abiotic stress (56 genes, one AP2 domain), ERF subfamily involved in biotic stress (65 genes, one AP2 domain), and others not yet classified (4 genes, one AP2 domain) (Sakuma et al. 2002). In this study, the DREB and ERF subfamilies were further divided into 12 subgroups suggesting that the family diverged in many functional differences. The exact number of AP2-related genes that exist in other plant species is unknown but based on bioinformatics searches plant ESTs, it appears probable that most plants will have representatives of each subfamily. The characterization of this gene family is still in its early stages. For example, 6 of the 12 subgroups in *Arabidopsis* are not yet characterized (Figure 2).

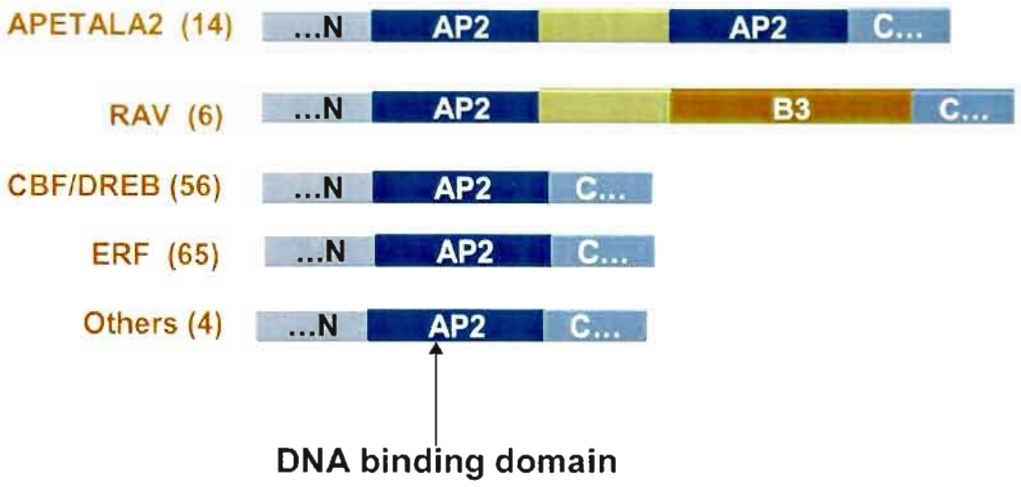


Figure 2: Structural organization of the AP2 transcription factors (according to Sakuma et al 2002).

3.2.1 Function and regulation of the AP2 subfamily

Members of the AP2 subfamily participate in the control of several steps or processes of flower development like organ and meristem identity, and ovule and seed development (Jofuku et al., 1994). In *Arabidopsis*, the homeotic gene *APETALA2* has been shown to control three important processes during flower development: (i) the establishment of flower meristem identity, (ii) the specification of flower organ identity and the regulation of floral organogenesis and (iii) the temporal and spatial regulation of flower homeotic gene activity (Riechmann and Meyerowitz 1998). This gene is involved in the specification of sepal and petal identity through its activity as a homeotic gene that forms part of the combinatorial genetic mechanism of floral organ

identity determination and in the negative regulation of the expression of the MADS-box floral homeotic gene AGAMOUS. *APETALA2* is expressed in the inflorescence meristem and throughout the floral primordia during the early stages of flower development; late in flower development, *APETALA2* transcripts are detectable in all organs but appear concentrated in petals and specific tissues of stamens and carpels, including ovules. In addition, *APETALA2* is expressed in vegetative leaves and in the stem (Jufuku et al., 1994). Keck et al (2003) used a reverse genetic approach, to show that the previous inability to obtain *Antirrhinum* mutants corresponding to the A class gene *APETALA2* of *Arabidopsis* reflects greater genetic redundancy in *Antirrhinum*. *Antirrhinum* has two genes corresponding to *APETALA2*, termed *LIP1* and *LIP2*, both of which need to be inactivated to give a mutant phenotype. Analysis of interactions between *LIP* and class *B/C* genes shows that unlike *APETALA2* in *Arabidopsis*, *LIP* genes are not required for repression of C in the outer whorls of the flower. *Arabidopsis* AINTEGUMENTA (ANT; Riechmann and Meyerowitz 1998) is required for ovule development and it also plays a role in floral organ growth, although it does not appear to have homeotic or organ identity functions. Maize Glossy 15 (Gl15; Riechmann and Meyerowitz 1998) regulates leaf epidermal cell identity. No members of this subfamily have been shown to be regulated by biotic or abiotic stresses (Sakuma et al 2002).

3.2.2 Function and regulation of the DREB and ERF subfamilies

The DREB/CBF and ERF subfamilies include 125 genes that have only 1 ERF/AP2 DNA-binding domain. From these, 121 contain a conserved WLC motif in the middle of their ERF/AP2 domains (Sakuma et al 2002). These 121 genes were further classified into 2 subfamilies based on similarities of the amino acid sequences

of their DNA-binding domain; 56 of them encode DREB/CBF-like proteins (group A) and 65 of them encode ERF-like proteins (group B). In addition to the conserved AP2 DNA-binding domain, AP2/EREBP protein shares two other features that are characteristic of transcription factors: regions of biased amino acid composition typical of transcription activation domains, and possible nuclear localization signals.

3.2.3 The DREB subfamily (group A)

Group A genes can be divided into 6 small subgroups based on similarities in the binding domain. The 1st and 2nd subgroups (A-1, A-2) include the *DREB1/CBF* and *DREB2* gene families, respectively, whose products bind to the DRE/CRT sequence. The 3rd subgroup (A-3) has only *ABI4*. The 4th subgroup (A-4) contains 16 genes, including *TINY*. The 5th subgroup (A-5) consists of 16 genes, including *RAP2.1*, *RAP2.9*, and *RA2.10*, and the 6th subgroup (A-6) contains 9 genes, including *RAP2.4*. DNA-binding specificities of the A-3, A-4, A-5, and A-6 genes remain unknown. *Arabidopsis* genome encodes a small family of cold-responsive transcriptional activators known either as CBF1, CBF2 and CBF3 (Gilmour et al., 1998) or DREB1b, DREB1c and DREB1a (Kasuga et al., 1999; Liu et al., 1998), respectively. These transcription factors recognize the cold and dehydration responsive DNA regulatory element CRT/DRE. These elements are present in the promoter regions of many cold and dehydration-responsive genes of *Arabidopsis*, including those designated COR (cold-regulated) (Thomashow 1999). The CBF genes are induced within 15 min of plants exposure to low, nonfreezing temperatures followed at 2 h by induction of cold-regulated genes that contain the CRT/DRE regulatory element i.e., the CBF regulon (Gilmour et al., 1998; Liu et al., 1998). Constitutive expression of the CBF genes in transgenic *Arabidopsis* plants results in the induction of *COR* gene expression and an increase in freezing tolerance without a low temperature stimulus (Gilmour et al., 2000; Kasuga et al., 1999). Significantly,

multiple biochemical changes that are associated with cold acclimation and are thought to contribute to increased freezing tolerance, including the accumulation of sugars and proline, occur in nonacclimated transgenic *Arabidopsis* plants that constitutively express CBF3 (Gilmour et al., 2000). Also, Gilmour et al (2000) proposed that the CBF genes act to integrate the activation of multiple components of the cold acclimation response. The results from Shinozaki et al (2003) indicated that activation of the CBF cold response pathway also results in enhancement of plant tolerance to drought and salt. Jaglo et al (2001) characterized two CBF-like proteins in *Brassica*. They found that the transcript levels for the *B. napus* CBF-like genes accumulate rapidly upon exposing plants to low temperature. The constitutive expression of *Arabidopsis* CBF2 and CBF3 in transgenic *B. napus* activates the expression of homologs of the CRT/DRE regulated *Arabidopsis* COR genes. In particular, the transcript levels for *Bn115* and *Bn28,* homologs of *Arabidopsis COR15a* and *COR6.6* respectively, were elevated in nonacclimated transgenic *B. napus* plants that overexpressed *CBF1*, *CBF2* or *CBF3*. Finally, expression of *Arabidopsis CBF* genes in transgenic *B. napus* resulted in an increased freezing tolerance. Jaglo et al (2001) identified a clone from tomato encoding a protein sharing significant sequence identity with *Arabidopsis CBF1*. Moreover, they found that *CBF*-like transcripts accumulate rapidly upon exposure of tomato plants to low temperature. The transcript levels of the tomato CBF in Castle Mart and other varieties appeared to return to those found in warm grown plants after 24 h of exposure to low temperature and remained at low levels after one week of cold treatment (Jaglo et al 2001). Hsieh et al (2002a) transformed a DNA cassette containing the *Arabidopsis CBF1* into the tomato genome. The transgenic tomato plants showed enhanced resistance to water deficit stress. Exogenous gibberellic acid treatment reversed the growth retardation and enhanced the growth of transgenic tomato plants, but did not affect the level of water deficit resistance. The stomata of the transgenic *CBF1* tomato plants closed more rapidly than the wild type after water deficit treatment with or without gibberellic acid pretreatment. The transgenic tomato

plants contained higher levels of proline than those of the wild-type plants under normal or water deficit conditions. Subtractive hybridization was used to isolate the responsive genes to heterologous *CBF1* in transgenic tomato plants and the isolated *CAT1* gene (*CATALASE1*) was characterized. Catalase activity increased, and hydrogen peroxide concentration decreased in transgenic tomato plants compared with the wild-type plants with or without water deficit stress. These results indicated that the heterologous *Arabidopsis CBF1* could confer water deficit resistance in transgenic tomato plants. The next question is whether species that have different cold sensitivity, have *CBF*-like genes that are rapidly induced in response to low temperature. Jaglo et al (2001) identified cDNA inserts encoding one wheat and three rye CBF-like polypeptides. As in *Arabidopsis* and *B. napus*, CBF-like transcripts accumulated rapidly in response to low temperature in both rye and wheat. Dubouzet et al (2003) isolated five cDNAs for *DREB* homologs from rice. Expression of *OsDREB1A* and *OsDREB1B* were induced by cold, whereas expression of *OsDREB2A* was induced by dehydration and high salt stresses. Overexpression of *OsDREB1A* in transgenic *Arabidopsis* induced the expression of target *COR* genes of *Arabidopsis DREB1A* resulting in plants with higher tolerance to drought, high-salt, and freezing stresses (Dubouzet et al 2003). This indicated that *OsDREB1A* has functional similarity to *DREB1A*. However, in microarray and RNA blot analyses, some *COR* target genes of the *Arabidopsis* DREB1A proteins that only have the DRE element ACCGAC were not over-expressed in the *OsDREB1A* transgenic *Arabidopsis*. The OsDREB1A protein bound to GCCGAC more preferentially than to ACCGAC whereas the DREB1A proteins bound to both GCCGAC and ACCGAC efficiently. The structures of DREB1 type ERF/AP2 domains in monocots are closely related to each other as compared with that in the dicots. Xue (2003) isolated HvCBF2 from barley leaves. In contrast to previously reported cold-inducible *CBF/DREB1* genes, *HvCBF2* was expressed in barley leaves under non-stress conditions. Two conserved amino acids in the ERF/AP2 domains differ between DREB and ERF. The valine at position 14 (V14) and the glutamic acid (E19) are conserved in the DREB proteins, whereas alanine and

aspartic acid are conserved at the corresponding positions of the ERF proteins. Sakuma et al (2002) confirmed the importance of these two conserved amino acid in determining the target DNA sequences in *Arabidopsis*.

3.2.4 The ERF subfamily (group B)

Group B genes encoding the ERF-type proteins can also be divided into 6 small subgroups based on the sequence identities of the DNA-binding domains. The 1st, 2nd, and 3rd subgroups (B-1, B-2, and B-3) include genes encoding proteins that bind specifically to the GCC box, such as ERF1, *AtEBP*, *AtERF-1*, *AtERF-2*, *AtERF-3*, *AtERF-4*, and *AtERF-5*. Therefore, gene products of these 3 subgroups probably bind to the GCC box or related sequences. The nucleotide sequences of *AtEBP*, *AtCdinp*, and *EBP*, which constitute the B-2 subgroup, are highly conserved and correspond to one genomic sequence (AB022217). There is no information concerning the DNA – binding sequence of proteins encoded by the ERF-related genes of subgroups B-4, B-5, and B-6. ERF proteins bind to a *cis* regulatory sequence widely conserved among ethylene-responsive pathogenesis-related (PR) genes, which are part of the battery of defense genes activated upon plant pathogen attack. Park et al (2001) reported that a tobacco gene encoding an ERF/AP2 transcription factor, *Tsi1*, was induced not only in leaves treated with ethephon, but also in leaves treated with NaCl or salicylic acid. The protein encoded by *Tsi1* could bind not only to the GCC box but also weakly to DRE. Overexpression of *Tsi1* improved tolerance to salt and pathogens. EREBP proteins may therefore form part of the mechanism used by plants to respond to biotic stress.

3.2.5 RAV subfamily

The 3rd class includes 6 genes that contain 2 different DNA-binding domains, ERF/AP2 and B3 (Kagaya et al 1999). The B3 DNA-binding domain is conserved in VP1/ABI3 homologues (Suzuki et al 1997). The B3 DNA-binding domain is shared amongst various plant-specific transcription factors, including factors involved in auxin-regulated and abscisic acid–regulated transcription.

The role of this group is not well understood but recently, the involvement of members of the RAV family in ethylene response (Alonso et al., 2003) and in brassinosteroid response (Hu et al., 2004) was reported. This subfamily is homologous to genes *AB13/VP1* found in maize. These genes containing VP1 domain act as intermediaries in regulating abscisic acid (ABA)-responsive genes during seed development. In rice, 4 members have been identified to date. The function and the target of proteins of subfamily RAV remain unknown. Two proteins, RAV1 and RAV2 were a subject for many researches. Gel shift assays were used for studies of binding and made it possible to see that RAV1 specifically bound to two motifs 5' - CAACA-3' and 5' - CACCTG-3' (Kagaya et al 1999).

3.2.6 Other AP2/EREBP members

The undefined subfamily comprises genes containing one or two AP2 domains which do not have any homology with the other subfamilies classified in this family (Sakuma et al 2002).

Recently, *AP2/ERF* domain-encoding genes were reported in bacteria, a bacteriophage, and a ciliate genome as a part of house-keeping endonuclease genes, mobile genetic elements that replicate and move in the genome (Magnani et al., 2004). In this report, it was also demonstrated that an AP2/ERF domain in a cyanobacterium, *Trichodesmium erythraeum*, recognizes stretches of poly (G)/poly(C), and that an

Arabidopsis ERF protein, AtERF#060 (At4g39780), contains a region similar to the HNH domain in the cyanobacterium AP2/ERF protein (Magnani et al., 2004).

4. Engineering transgenic abiotic stress-tolerant plants with transcription factors

Several key transcription factors were identified under abiotic stress treatments based on molecular and genomic studies. Among them *DREB* and *ABF* are well characterized transcription factors known to play an important role in regulating gene expression in response to abiotic stresses via ABA-independent and ABA-dependent manner. Kasuga et al (1999) overexpressed the cDNA encoding *DREB1a* under the control of the CaMV35S promoter in transgenic *Arabidopsis* plants. As a result, many of stress tolerance genes such as *COR*, *ERD*, *P5CS* and *rd29* are expressed under normal growing conditions and likely explain the improved tolerance to drought, salinity, and freezing stress. However, constitutive expression of *DREB1a* resulted in severe growth retardation under normal growing conditions. In contrast, expression of *DREB1a* gene under the control of a stress-inducible *rd29A* promoter gave rise to minimal effects on plant growth under normal growing conditions and provided even greater tolerance to abiotic stress treatments. Oh et al (2005) ectopically expressed *Arabidopsis DREB1a* (*CBF3*) in transgenic rice plants under the control of constitutive promoter. The authors reported that *DREB1a* transgenic rice plants show enhanced tolerance to drought and salinity but limited low temperature tolerance, with no visible stunted phenotype despite its constitutive expression. These results differ from the *Arabidopsis* results in relation to freezing tolerance as reported earlier from Kasuga et al (1999). This could be partly explained by the finding expression of different sets of stress-related genes such as *Lip5*, *Dip1*, *PSLS*, *HSP70*, *PP2Ca*, *Rab21* in *DREB1a* transgenic rice plants (Oh et al 2005), which are known to enhance osmotic stress tolerance. In another study, Ito et al (2006) analyzed *OsDREB1* transgenic rice plants showing improved tolerance to drought, salt and low temperatures and identified large number of stress-inducible genes.

These results confirm that the DREB1/CBF cold-responsive pathway is conserved in rice and *Arabidopsis*. Overexpression of *CBF1* (*DREB1b*) in *Arabidopsis* and *Brassica* induces *COR* genes in both species (Jaglo et al 2001) but not in tomato (Hsieh et al 2002a). Hsieh et al (2002b) reported that drought, chilling and oxidative stress tolerance were improved in tomato plants expressing *Arabidopsis CBF1*.

The function of transcription factors in response to ABA was also identified using transgenic approaches. ABA-responsive element binding factor (*ABF*) members, which belong to the bZIP transcription factor family, show distinct roles in response to ABA and other abiotic stresses (Uno et al 2000). To reveal how *ABFs* are involved in stress tolerance, Kang et al. (2002) generated *ABF3* and *ABF4* transgenic *Arabidopsis* lines by overexpressing them constitutively. Both transgenic lines exhibited an altered transpiration rate in response to water deficit conditions. All transgenic plants survived a 12-day drought treatment and set seed in contrast to wild type plants with 33% survival rate. Both overexpression lines also showed induction of ABA-signaling genes including *ABI1*, *ABI2* phosphatase, and other stress-responsive genes including desiccation-related *LEA* genes via an ABA-dependent pathway (Kang et al 2002). Both transgenic lines were also hypersensitive to salinity treatments during germination. It is interesting that although both *ABF3*and *ABF4* play an essential role in drought tolerance, their constitutive overexpression resulted in a stunted phenotype in the *ABF4* line and no visible growth inhibition in *ABF3* lines in comparison to wild type plants (Kang et al 2002). Similarly, constitutive overexpression of *ABF3* in rice did not show visible phenotypic growth alterations and plants drought tolerance with activation of stress-regulated genes, *Wsi18* and *Rab21* (Oh et al 2005). Kim et al. (2004) generated *ABF2* transgenic *Arabidopsis* plants under the control of 35S promoter and the overexpression lines exhibited seedling growth inhibition. The phenotype was derived from the repression of hexokinase *HXK1* and *HXK2* expression that was eventually relieved by the addition of sucrose, suggesting that *ABF2* is involved in the sucrose response pathway. In contrast to *ABF3* and *ABF4*, the two-week old *ABF2* transgenic plants showed only

10–15% survival during drought stress and exhibited higher salt tolerance. However, when 3-week old *ABF2* transgenic plants were tested for drought tolerance, nearly 97% of the transgenic plants survived in comparison to only 22% survival rate seen in wild type plants (Kim et al 2004). So far, *ABF1* overexpression or knockout phenotypes were not described in *Arabidopsis*. Recently, Furihata et al (2006) demonstrated that overexpression of the *Arabidopsis* phosphorylated active form of AREB1 (*ABF2*) resulted not only in the induction of the stress-responsive *RD29B* gene but also of seed-specific genes without ABA treatment. Among the bZIP family, constitutive expression of a pathogen-induced pepper bZIP transcription factor (*CAbZIP1*) in transgenic plants exhibited resistance to the pathogen *Pseudomonas syringae*, and drought and salt tolerance during all stages of plant growth (Lee et al 2006).

One important way of achieving tolerance to multiple stress conditions is to overexpress transcription factor genes that control multiple genes from various pathways. Although an important role of the well characterized *DREB* and *ABF* stress-responsive transcription factors was revealed recently in conferring abiotic stress tolerance by transgenic approaches, very little is known about the function of other transcription factors that act under the influence of phytohormones such as ethylene, jasmonic acid, and salicylic acid, which confer abiotic stress tolerance. The ethylene-responsive element binding proteins (EREBP) belong to the AP2 transcription factor family, which is known to act under the influence of ethylene that mediates plant responses to both biotic and abiotic stresses. Park et al (2001) identified a stress-induced EREBP transcription factor (*Tsi1*), known to bind GCC and DRE/CRT boxes in tobacco, which was induced by ethylene, methyl jasmonate, salicylic acid, salt treatments and wounding. Overexpressing transgenic tobacco lines showed enhanced salt tolerance and resistance to pathogens (Park et al., 2001). In another study, the ethylene response factor (*ERF*), a member of the EREBP transcription family from tomato known to be responsive to jasmonate and ethylene, was constitutively overexpressed in transgenic tobacco and found to confer enhanced

tolerance to salt and pathogens by activating expression of pathogen-related genes (Wang et al 2004). Based on these results, it is apparent that some of the stress-responsive EREBP family members are potential regulators in ethylene and osmotic stress signaling.

RESEARCH GOAL

Cereal crops constitute more than 60% of total worldwide agricultural production. The most important cereal crops are rice, wheat and maize. More than 500 millions tons of each is produced annually worldwide (Harlan 1995). The world food grain production needs to be doubled by the year 2050 to meet the ever growing demands of the population (Tilman et al., 2002). This goal needs to be achieved despite decreased arable land, declining water resources, and the environmental constraints such as drought, excess heat, frost, salinity, metal toxicity and nutrient imbalances, which cause major losses in cereal grain production. In response to biotic and abiotic stresses, plants modulate the expression of specific sets of genes to cope with these stresses (Shinozaki and Yamaguchi-Shinozaki 1997). Stresses such as salinity, pathogen attack, wounding, UV irradiation, high or low temperature, drought, and the hormones ethylene and ABA have been reported to induce genes that encode pathogenesis-related (PR), antifungal proteins (Chakravarthy et al 2003), late embryogenesis abundant (LEA) (NDong et al 2002) and cold-regulated (COR) proteins (Takumi et al 2003). Although these adaptations are adequate to sustain the plants survival in its natural habitat, they are at present becoming limiting because the increased human population and industrial activity is forcing crop cultivation to areas less fit for plant growth. Extensive genetic studies and surveys of landrace and wild germplasms have indicated extensive variation for abiotic stress tolerance but this has been difficult to exploit due to the relatively poor background knowledge of the molecular basis for stress in these species (Langridge et al 2006). Interconnected signal transduction pathways that lead to multiple responses to abiotic stresses have been difficult to study using traditional approaches because of their complexity and the large number of genes and gene products involved in the various defensive and

developmental responses of the plant. Therefore solutions to increase the plant resistance to these stresses must be found (Langridge et al., 2006).

Functional genomics is now widely used as tools for dissecting abiotic stress responses in cereals, through which networks of stress perception, signal transduction and defensive responses can be examined from gene transcription, through protein complements of cells, to the metabolite profiles of stressed tissues. Another way is to develop crops that are more tolerant to abiotic stresses, and this will allow growth on lands previously not fit for cultivation. Past efforts to improve plant tolerance to drought, high salinity and low-temperature through breeding and genetic engineering have had limited success owing to the genetic complexity of stress responses. However, recent advances have shown that a better understanding of a transcription regulation may be useful in simplifying the genetic engineering of stress tolerance in plants. For example, the transcription factors ethylene-responsive element binding proteins (EREBPs) and ERFs containing the APETALA2 (AP2)/ERF- DNA binding domain have been shown to play crucial roles in various types of biotic and environmental stresses. A major breakthrough in improving a complex trait such as plant stress tolerance was reported by Jaglo-Ottosen et al (1998). They found that overexpression of the CBF/DREB1 proteins in *Arabidopsis* resulted in increased freezing tolerance at the whole plant level in both nonacclimated and cold-acclimated transgenic plants. Other transgenic studies confirmed this finding with members of the DREB and ERF subfamilies. Overexpression of ERF1 resulted in the induction of several ethylene responsive defense genes in *Arabidopsis* and conferred resistance to necrotrophic fungi such as *B. cinerea* and *Plectosphaerella cucumerina* (Stepanova and Ecker 2000). These results showed the importance of AP2 family as a key regulator for several abiotic and biotic stresses and the impact to use some members from this family to improve the tolerance of wheat plants. In spite of all these discoveries, less is known about the regulation of AP2 in cereals in comparison with *Arabidopsis*. The goal of this project is to identify and characterize the LT-regulated

members of the AP2 family. The functional analysis of the key members may be used as a tool to improve stress tolerance of wheat and other cereals to different stresses.

OBJECTIVES

Abiotic and biotic stresses cause major losses in crop productivity worldwide. For this reason, a great deal of research has been carried out in the past few years to try to understand the nature of the genetic regulatory systems involved. Among the components involved in these regulatory pathways is the plant specific transcription factor superfamily AP2 (for APETALA2). This protein family comprises 144 members in *Arabidopsis* (Sakuma et al 2002), has in common the AP2 DNA-binding motif. Members of this family were shown to be regulated by cold, drought, pathogen infection, salinity, wounding or treatment with ethylene, salicylic acid or jasmonic acid. Direct evidence has accumulated that members of this family can have a major impact on complex plant stress responses. An important question that remains to be answered is whether the wide spectrum of plant adaptive responses to stresses is mediated, in part, through the diversification of the structure and/or the regulation of this transcription factor family. Phylogenetic analysis of the AP2 gene family, using *Arabidopsis* and rice as model genomes, offers a useful tool with which to approach this question. In the first article, we present an overview of our efforts to identify and classify expressed *AP2* genes from wheat. The functional characterization of this family in cereal species such as wheat, an economically important crop in Canada and the world, will have an important impact on the development of appropriate strategies to manipulate this transcription factor family for producing stress tolerant plants.

It has recently been established that an important component of cold acclimation in *Arabidopsis* is the CBF (C-repeat Binding Factors) cold response pathway (Jaglo et al 2001). Within 15 min of exposing plants to low temperature, genes that encode a small family of transcriptional activators known as CBF1, CBF2, and CBF3 are induced. The CBF proteins belong to the AP2 family of DNA-binding proteins. Direct evidence has accumulated that members of the CBF subgroup can

have a major impact on complex plant stress responses. Expression of the CBF regulon results in an increase in freezing tolerance due, in part, to the induction of genes encoding cryoprotective proteins and enzymes involved in the synthesis of low-molecular weight cryoprotectants. An important question that remains to be answered is whether the wide spectrum of plant adaptive responses to stresses is mediated, in part, through the diversification of the structure and/or the regulation of this transcription factor family. Phylogenetic analysis of the CBF gene family, using *Arabidopsis* and rice as model genomes, offers a useful tool with which to approach this question. In the second article, I present an overview of my effort to identify and classify expressed *CBF* genes from wheat. The functional characterization of this subgroup in economically important cereals such as wheat will facilitate the production of stress tolerant plants.

In comparison with the dicot *Arabidopsis*, monocot plants, whether freezing tolerant or not, seem to have a larger *CBF* gene family. In particular, freezing tolerant plants such as wheat and barley have a rather large and complex CBF family containing more than 25 genes (Badawi et al. 2007, Skinner et al. 2005). The expansion of the CBF family suggests that CBF genes may play important functions other than cold response and raises a question whether several *ICE* genes are involved in the regulation of this big family in wheat. The third article will demonstrate that there are at least two different *TaICE* genes able to activate wheat or *Arabidopsis* CBF promoters.

ARTICLE I

IDENTIFICATION AND CLASSIFICATION OF THE WHEAT AP2/EREBP TRANSCRIPTION FACTOR GENE FAMILY

Mohamed Badawi, Jean Danyluk, Barbara Boucho and Fathey Sarhan*

Département des Sciences Biologiques, Université du Québec à Montréal, Case Postale 8888, Succursale Centre-ville, Montréal, Québec, Canada H3C 3P8

Footnotes

Financial source: This work was supported by a Natural Sciences and Engineering Research Council of Canada discovery grant (F.S.), and by Genome Canada and Génome Québec

Key words: Hexaploid wheat, AP2 protein, biotic stress, abiotic stress.

Running Title: "AP2 genes in hexaploid wheat".

* To whom correspondence should be addressed.

Contribution

I took part in all the stages of experimentation under the supervision of FS and JD. I was involved in the identification, the cloning and the sequence analyses of the *AP2* genes with BB and JD. I performed the bioinformatic and phylogenetic analyses. I also took an active part in drafting the manuscript.

Résumé

Les céréales sont exposées à plusieurs stress biotiques et abiotiques durant leur cycle de vie. Un déficit en eau suite à des conditions de sécheresse, de basses températures ou de salinité, constitue l'un des stress les plus limitants pour la croissance et le développement des plantes. Ces différentes conditions de stress peuvent affecter le métabolisme des plantes ainsi que l'expression de nombreux gènes. La famille de gènes qui codent pour des facteurs de transcription AP2/EREBP est particulièrement ciblée lors de stress environnementaux. Des travaux ont montré que l'expression de gènes *AP2/EREBP*, en réponse à plusieurs types de stress, était modulée pour contrôler divers processus d'adaptation et du développement de la plante. Une recherche dans les bases de données publiques NCBI a permis de reconstituer 133 séquences contiguës (contigs) de blé codant pour différents facteurs AP2/EREBP. Un criblage de banques d'ADN complémentaires (ADNc), les ADNc générés par le projet FGAS (Functional Genomics of Abiotic Stress) et l'amplification par réaction de polymérisation en chaîne (PCR) ont permis l'identification et le clonage de 107 *AP2/EREBP* gènes de blé. L'organisation structurale de ces gènes du blé ainsi que leur relation phylogénétique avec leurs homologues chez *Arabidopsis* ont montré qu'en dépit d'une forte conservation de fonction chez les angiospermes, quelques divergences et éléments structuraux sont frappants entre les espèces monocotylédones et dicotylédones. Une compréhension de l'évolution de la famille *AP2/EREBP* et leur caractérisation fonctionnelle chez les espèces céréalières économiquement importantes (blé, orge, riz et maïs) peuvent être utiles pour développer des espèces plus tolérantes aux stress ou des stratégies de cultures appropriées.

Mots clés: blé, facteur de transcription, AP2/ERF, le stress, L'évolution, à basse température

Abstract

Crop plants are exposed to many types of abiotic and biotic stresses during their life cycle. Water deficit resulting from drought, low temperature or high salt concentration in the soil, is one of the most common environmental stresses that affect growth and development in plants. These stresses can alter plant metabolism and gene expression. Among the regulatory genes modulated by different stresses is the plant specific *AP2/EREBP* transcription factor gene family. Members of this superfamily have been shown to play a variety of roles as key regulators of several developmental processes and form part of the mechanisms used by plants to respond to various types of biotic and environmental stresses. We used data mining of NCBI databases and Cap3 assembly to generate 133 contigs corresponding to different *AP2* genes in wheat. Using cDNA library screening, the FGAS (Functional Genomics of Abiotic Stress) collection and PCR amplification, we have identified and fully sequenced 107 cDNAs from wheat. An overview of this gene family in wheat is presented, including the gene structures, phylogeny and chromosome locations. In addition, a comparative analysis between these genes in *Arabidopsis*, rice and wheat was performed. A phylogenetic overview of these genes suggests that despite broad conservation of their function in monocots and dicots, some structural elements are specialized within each of these three lineages. Understanding the evolution and the functional characterization of this family in economically important cereal species such as wheat, barley, rice and maize will have an important impact on the development of appropriate strategies to manipulate this transcription factor family for producing stress tolerant plants.

Key words: Wheat, transcription factor, AP2/ERF, stress, evolution, low temperature

Introduction

Abiotic and biotic stresses cause major losses in crop productivity worldwide. For this reason, a great deal of research has been carried out in the past few years to try to understand the nature of the genetic regulatory systems involved. A considerable number of genetic and molecular studies have demonstrated that stress-responsive signaling pathways in plants involve a complex network of various components, including receptors, kinases, phosphatases, and transcription factors (Gutterson and Reuber 2004). Transcription factors are believed to play a crucial role in the regulation of downstream gene expression of signal transduction pathways during abiotic and biotic stresses. In *Arabidopsis*, various transcription factor families, containing functional domains such as AP2, bZIP/HD-ZIP, WRKY, MYB, and several classes of zinc-fingers, have been studied in this context (Li et al 2004; Riechmann et al 2000). The AP2/EREBP superfamily is defined by the AP2/ERF domain, which consists of about 60 to 70 amino acids and is involved in DNA binding (AP2; Jofuku et al. 1994). The plant specific transcription factor superfamily AP2 family that includes 147 and 165 gene members in *Arabidopsis* and rice, respectively (for review see Riechmann et al 2000; Nakano et al 2006).

The proteins encoded by the *AP2/EREBP* gene family have diverse functions throughout the plant life cycle, including regulation of development, responses to abiotic stresses such as drought , cold, salinity, wounding as well as to biotic stresses such as fungal pathogen infections treatment with ethylene, salicylic acid or jasmonic acid (Feng et al 2005; Banno et al 2001; Dubouzet et al 2003; Ohto et al 2005). The AP2/EREBP family is divided into the RAV, AP2, and EREBP subfamilies, with the EREBP subfamily being divided into DREB or A subgroup and the ERF or B subgroup (Sakuma et al 2002). The AP2 family proteins contain two repeated AP2/ERF domains, the ERF family proteins contain a single AP2/ERF domain, and

the RAV family proteins contain single AP2/ERF domain and a B3 domain, which is a DNA-binding domain conserved in other plant-specific transcription factors, including VP1/ABI3, (Sakuma et al 2002). Group A genes contains 56 genes encoding DREB/CBF-like proteins that can be divided into 6 small subgroups based on similarities in the binding domain. The 1st and 2nd subgroups (A-1, A-2) include the *DREB1/ CBF* and *DREB2* gene families, respectively, whose products bind to the DRE/CRT sequence. The 3rd subgroup (A-3) has only *ABI4* which is expressed in seeds and regulates some seedling responses to ABA (Sakuma et al., 2002). The 4th subgroup (A-4) contains 16 genes, including *TINY*. The 5th subgroup (A-5) consists of 16 genes, including *RAP2.1, RAP2.9,* and *RAP2.10,* and the 6th subgroup (A-6) contains 9 genes, including *RAP2.4*. The ERF or B subgroup contains 65 *ERF* genes and contains all of the *AP2/EREBP* genes that have been linked to disease resistance responses (Gutterson and Reuber, 2004). The AP2 domain has been considered plant specific (Riechmann and Meyerowitz 1998). However, recent studies showed that homologues are present in the cyanobacterium *Trichodesmium erythraeum*, the ciliate *Tetrahymena thermophila*, and the viruses *Enterobacteria phage Rb49* and *Bacteriophage Felix 01* (Magnani et al 2004; Wuitschick et al 2004; Wessler 2005). These nonplant proteins bearing an AP2 domain are predicted to be HNH (or in some cases, HNN; histidine and asparagine) endonucleases. Magnani et al (2004) hypothesized that a horizontal transfer of an HNH-AP2 endonuclease from bacteria into plants may have led to the origin of the *AP2/EREBP* family.

The availability of the rice (*Oryza sativa*) genome sequences allowed a comparative analysis between *Arabidopsis* and rice within the AP2 family, which is useful in terms of studying the functional and evolutional diversity of the transcription factor family in plants. Determining the phylogenetic relationships of the *AP2/EREBP* multigene family among plants is an important step in elucidating the evolution of this developmentally and physiologically important gene family. An important question that remains to be answered is whether the wide spectrum of plant adaptive responses to stresses is mediated, in part, through the diversification of the

structure and/or the regulation of this transcription factor family. Phylogenetic analysis of the AP2 gene family, using *Arabidopsis* and rice as model genomes, offers a useful tool with which to approach this question. In wheat, some genes encoding putative AP2/EREBP proteins were identified. However, very few *AP2/EREBP* genes in wheat have known function. In the present study, we identified 107 *AP2/EREBP* genes from hexaploid wheat. Phylogenetic and molecular analyses indicated a broad conservation of *AP2/EREBP* gene sequences and functions between wheat and previously described models such rice and *Arabidopsis*. The CBF-like genes seemed to be evolved in monocots and probably achieving novel functions during wheat adaptability to LT. Some other *AP2/EREBP* genes were clearly classified into group A and B may have functions during drought and biotic stress in wheat.

Materials and methods

Plant materials and growth conditions

One spring habit (cv. Quantum) and one winter habit (cv. Norstar) hexaploid wheat cultivars (*Triticum aestivum* L. AABBDD x 7 = 42 chromosomes) were grown in environmentally-controlled growth chambers as previously described (Badawi et al 2007). Briefly, seedlings were germinated for one week under long day (LD, 16h) photoperiod at 20°C. For LT treatment, plants grown for 7 days at 20°C under LD photoperiod were transferred at 4°C under identical photoperiods.

Library construction and cDNA sequencing

The procedure of wheat cDNA libraries was described previously (Houde et al 2006; Badawi et al 2007). Full-length sequences of the cDNAs were obtained by sequencing from both the 5' and 3' ends using SM13F, M13R and T7 primers on the CEQ™ 2000 DNA Analysis System (Beckman). Internal primers were designed and used to complete sequencing, as needed (data not shown).

Data mining and genes isolation

BLAST searches of the NCBI, TIGR and FGAS (Functional Genomics of Abiotic Stress) databases for related wheat EST sequences were performed using the AP2 domains as queries. Sequences retrieved by this method were used as query to search for their closest relative in *Arabidopsis*, rice and wheat genomes. Wheat ESTs exhibiting similarities (> 40%) to different clades of *AP2* genes were assembled using the Cap3 sequence assembly program (http://deepc2.psi.iastate.edu/aat/cap /cap.html). The assembly produced 133 contigs (virtual mRNAs containing at least 2 ESTs) and over 50 singletons. Details of the screening and PCR isolation were described

elsewhere (Badawi et al 2007). This approach was used and proven to be efficient (Badawi et al 2007; Houde et al 2006).

Phylogenetic analyses

A phylogenetic reconstruction of AP2 family members was performed using the AP2 domain (60 amino acids) from 147 sequences of *Arabidopsis*, 151 of rice and 76 of wheat. Phylogenetic analysis of the 76 wheat AP2 proteins was performed using the AP2 domain. All alignments of AP2 amino acid sequences were performed using ClustalW with the following parameters: gap opening penalty of 10.00, gap extension penalty of 0.20 and BLOSUM protein weight matrix. To construct reliable trees, all have been manually corrected and resubmitted to multiple alignments. Except for wheat AP2 sequences, all the accession numbers and the alignments used to generate trees are available upon request. To better visualize the relationship between wheat *AP2* genes, only one copy of each gene was included in each analysis. If a copy was isolated from two different cultivars, only the one from Norstar was used in subsequent analysis.

Analysis of AP2 protein sequences

The analysis of shared motifs among the wheat AP2 protein sequences was performed using the MEME version 3.0 (Bailey and Elkan 1994) as previously reported (Parenicova et al 2003). The InterProScan (http://www.ebi.ac.uk/InterProScan/, Quevillon et al 2005) and the NCBI conserved domain search (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi, Marchler-Bauer et al, 2005) were also used to support the results obtained through the MEME analysis.

Chromosomal location

A series of cytogenetic stocks of the wheat cultivar Chinese Spring were obtained from the Wheat Genetics Resources Center, Kansas State University, USA

(Sears 1954, 1966; Endo and Gill 1996). The series included 19 nullisomic-tetrasomic lines, 23 ditelosomic lines and 96 deletion lines.

Leaf tissues was collected (5g) from four week old plants grown in growth chambers, freeze-dried and ground in 50 ml centrifuging tubes with the help of 2mm glass beads using a high speed shaker. Fifteen micrograms of genomic DNA from 138 wheat lines used in this study was digested with the *Dra I* restriction enzyme for 4 hours at 37°C in the presence of RNAse A (0.03 μg/reaction). For hybridizations, each probe was PCR-amplified and 100 ng purified DNA from each clone was labelled with α-^{32}p using a random primer labelling kit following the manufacturers instructions (Invitrogen) and purified using Sephadex G-50 columns (Amersham). The filters were hybridized, washed at high stringency and exposed to BioMax MS, X-Ray film through a BioMax TranScreen HE intensifying screen (Rochester, New York) for 1 to 10 days depending on the signal strength.

Results

Identification of *AP2* genes in hexaploid wheat

To initiate this project, available AP2 protein sequences were used for data mining of the NCBI databases in search of wheat homologs. ESTs from the wheat genomics of abiotic stress (WGAS) project (http://www.bioinfouqam.wgas.ca/cgi-bin/abiotic/project.cgi) were ordered and completely sequenced at the McGill University and Genome Quebec Innovation Center (Montreal, Canada). Wheat mRNAs and ESTs were assembled using the CAP3 sequence assembly program (http://deepc2.psi.iastate.edu/aat/cap/cap.html) into virtual mRNAs encoding putative AP2 proteins. These sequences were used as query to search for their closest relatives in *Arabidopsis* and rice genomes. To increase chances of identification of AP2 genes in wheat, we used a strategy combining data mining, cDNAs libraries screening and PCR amplification as previously described (Houde et al 2006, Badawi et al 2007). The results allowed the isolation of 107 complete cDNAs encoding putative AP2 (Table I). Based on the identity and homology at the nucleic acid level (97% percent identity and a 40 overlap cut-off), these expressed sequences represented 76 different AP2 genes (Table I).

Recently, 37 cDNAs encoding putative AP2 in hexaploid wheat were reported (Badawi et al 2007). A close analysis of these expressed sequences revealed that they correspond to 15 nonredundant genes. Together, this analysis indicated that at least 76 nonredundant genes are expressed in hexaploid wheat (Table I, Figure 1). The biological functions of wheat AP2 protein remain to be elucidated.

Phylogenetic Analysis of the *TaAP2* Genes

Sakuma et al (2002) and Nakano et al (2006) have provided the valuable overview of the single-AP2-domain-containing TFs. This group comprises approximately 125 genes that fall into two broad subclasses, those with conserved binding domains that are most similar to the dehydration-responsive element binding (DREB) genes (A subfamily) and those that are most similar to AtERF1 (B subfamily). Sakuma et al. 2002 also divided each subgroup into six further subfamilies on the basis of conserved regions. Our own analysis of the full *Arabidopsis AP2* gene family (comprising 147 members) is consistent with theirs, with the additional members falling into a subfamily whose members are most related to APETALA2 (which contains two AP2 domains) and a small subfamily whose members are related to RAV1 (and thus have a B3 domain in addition to the AP2 domain). Proteins belonging to the CBF/DREB1 transcription factor family have previously been shown to be the regulators of the majority of cold-response genes. In wheat, we found 107 genes belonging to the AP2/ERF superfamily (Table 1). To obtain clues about the evolutionary of the *TaAP2* genes in wheat, a phylogenetic tree was generated using the multiple sequence alignments of the *TaAP2* protein sequences. The resulting tree revealed that 53 of the wheat AP2/ERF proteins belong to the DREB subfamily, 3 to the AP2/RAV subfamily, 48 to ERF subfamily and the remaining 3 sequences to the APETALA2 subfamily (see Figure 1 and Table 1). From the analysis it can also be deduced that the wheat CBF/DREB1 proteins are most closely related to the monocot CBF/DREB1 proteins (Figure 2).

APETALA2-like subfamily

Genes in the AP2 subfamily function as key developmental regulators in reproductive and vegetative organs (Riechmann and Meyerowitz 1998). The AP2 subfamily is further divided into two monophyletic groups: AP2 and ANT (Shigyo and Ito 2004), including the floral homeotic gene *APETALA2*(*AP2*) (Jofuku et al

1994) and *AINTEGUMENTA* (*ANT*; Klucher et al 1996), which is involved in lateral organ development by controlling cell number and growth (Mizukami 2001). Three sequences were identified in wheat (TaANTL1B, TaAPETALA2, and TaBBM2).

RAV-like subfamily

The transcription factors RAV1 and RAV2 contain an AP2 domain in the N-terminal region, and a B3 domain in the C-terminal region (Kagaya et al 1999). No plant transcription factor has yet been demonstrated to contain two or more DNA-binding domains of distinct types, except for RAV1 and RAV2 (RAV: Related to ABI3/VP1) from *Arabidopsis* (Kagaya et al 1999). Using binding site selection assays, the AP2 and B3 domains of RAV1 were found to bind to the CAACA and CACCTG motifs respectively (Kagaya et al 1999). RAV1 functions as a transcriptional activator triggering resistance to bacterial infection and tolerance to osmotic stresses in pepper (Sohn et al 2006). Three genes were identified in wheat (TaRAV1-1, TaRAV1-2 and TaRAV2).

DREB-like subfamily

In *Arabidopsis* this subfamily contains 56 genes, encoding the DREB/CBF-like proteins and can be divided into 6 small subgroups based on similarities in the binding domain. The 1st and 2nd subgroups (A-1, A-2) include the *DREB1/CBF* and *DREB2* gene families, respectively, whose products bind to the DRE/CRT sequence. The 3rd subgroup (A-3) has only *ABI4*. The 4th subgroup (A-4) contains 16 genes, including *TINY*. The 5th subgroup (A-5) consists of 16 genes, including *RAP2.1, RAP2.9,* and *RAP2.10,* and the 6th subgroup (A-6) contains 9 genes, including *RAP2.4*. The dehydration responsive-element binding proteins (DREB) are important transcription factors that induce a set of abiotic stress-related genes and impart stress endurance to plants. The DREB transcription factors could be dichotomized as DREB1 and DREB2, which are involved in two separate signal transduction pathways under low temperature and dehydration, respectively. In wheat, 53 genes

from this subfamily were isloted under our experimental conditions and classified as follows: (A-1) 37 genes; (A-2) 4 genes; (A-3) 1 gene; (A-4) 2 genes; (A-5) 7 genes; (A-6) 2 genes. It is difficult to assess the general picture of DREB regulation in relation to phylogeny because different studies have concentrated only on the CBF group and DREB2A.

EREBP-like subfamily

The ERF-type proteins are divided into 6 small subgroups based on the sequence identities of the DNA-binding domains. The 1st, 2nd, and 3rd subgroups (B-1, B-2, and B-3) include genes encoding proteins that bind specifically to the GCC box, such as *ERF1, AtEBP, AtERF-1, AtERF-2, AtERF-3, AtERF-4,* and *AtERF-5.* There is no information concerning the DNA-binding sequence of proteins encoded by the ERF-related genes of the subgroups B-4, B-5, and B-6. The proteins encoded by *AC016972, RAP2.11, AC006258, AC360314,* and *AC007591* in subgroups B5 and B6 have a valine at the AA14 position. ERF genes show a variety of stress-regulated expression patterns. Regulation by disease-related stimuli such as ethylene (ET), jasmonic acid (JA), salicylic acid (SA) and infection by virulent or avirulent pathogens, has been shown for several ERF genes (Brown et al 2003, Fujimoto et al 2000). However, some ERF genes are also induced by wounding and abiotic stresses (Park et al 2001). In wheat, 48 genes from this subfamily were isolated under our experimental conditions and classfied as follows: (B-1) 23 genes; (B-2) 10 genes; (B-3) 10 genes; (B-4) 1 gene; (B-5) 2 genes; (B-6) 2 genes. At present, it is difficult to assess the overall picture of ERF regulation in relation to phylogeny because different studies have concentrated on different ERF genes, treatments and time points.

Chromosome mapping

The map position of AP2 genes in wheat genomes was used to discriminate between homeologues (copies of genes) as well to discover whether any of them could be linked to known QTLs having an effect on FT. The map position of several AP2 sequences was determined by Southern blot analysis. Specific probes and a series of cytogenetic stocks of the wheat cultivar Chinese Spring were used to map genes to bins corresponding to chromosomal deletion regions. The results presented in Table 3 indicate that members of the family are scattered through the 7 chromosomes of all three genomes of wheat. There is a particular clustering of AP2 gene mapped on the long arm of group 5 chromosomes (Table 3). In *Triticeae* species, this region is linked to the frost tolerance locus *Fr-A* ^{m} (*2*). Similarly, in the barley study, QTLs for frost tolerance were detected at the *VRN-H1* and the *Fr-H2* loci, but only the QTL for *Fr-H2* overlapped a QTL for differential accumulation of COR14b protein in leaf samples collected from the field at the beginning of the winter (Francia et al 2004). Some mapping results presented are consistent with some previously reported data (Badawi et al. 2007).

Discussion

AP2 genes in hexaploid wheat

Cloning and characterization of transcription factors family genes has contributed greatly to our understanding of the physiological responses of plant cells at the molecular level to different environmental stresses and development. Among most important transcription factor impacting plants are AP2s. Because of their highly conservation among eukaryotes, plant *AP2* genes are excellent in the elucidation of developmental and environmental stresses processes, in both phylogenetic and genetic features.

Bread wheat is a hexaploid species with genome constitution AABBDD, which originated from three diploid relative species: A genome from *T. urartu*, B genome from *Aegilops speltoides* or other species classified into Sitopsis section, and D genome from *Ae. tauschii* (Feldman 2001). Allopolyploidization leads to the generation of duplicated homoeologous genes, which is opposed to paralogous genes. Consequently, the hexaploid wheat genome contains triplicated homoeologous genes derived from the ancestral diploid species. For geneticists and molecular biologists, the first challenge using such a model is to differentiate copies of genes and duplication events during the genes evolution-development.

In this study, comparative approaches were used to distinguish between wheat *AP2* homeologous genes copies. All sequences analysed were from hexaploid wheat (cv. Norstar) and then were treated arbitrary as duplicates if they were > 95% identical at the DNA level with 40 overlap cut-off. Sequence analysis led to the recognition of 107 nonredundant AP2 genes. On the other hand, additional virtual cDNAs were reconstructed from initial data mining but failed to be amplified under

our experimental conditions. Therefore, it is apparent that the wheat genomes encode more AP2 proteins than the 107 members identified so far.

Chromosomal location of the genes was used to discriminate between copies. Although *AP2* genes are scattered on the wheat genomes, some closely related genes are mapped on different group chromosomes. This finding is consistent with previous data reported for *Arabidopsis*; 122 *ERF* genes are distributed over all of the chromosomes. *TaAPETALA2* and *TaANTL1B* genes were located on different chromosomes, 5 and 7, respectively in wheat. *CBF* genes are clustered and mapped in the long arm of group chromosomes 5. Many *TaCBF* genes are located on two or more distinct genomes possibly because re-arrangements of chromosomes occurred between genomes.

AP2 genes classification and phylogeny

To determine the phylogenetic relationships among the AP2 gene family in wheat, a multiple alignment analysis was performed using amino acid sequences in the AP2/ERF domain. A comprehensive overview of the single-AP2-domain-containing AP2 TFs was reported by Sakuma et al (2002) and Nakano et al (2006). This group comprises approximately 125 and 139 genes in *Arabidopsis* and rice, respectively. These genes fall into two broad subclasses, those with conserved binding domains are most similar to the dehydration-responsive element binding (DREB) genes (A subfamily) and those that are most similar to AtERF1 (B subfamily). Our own analysis of the wheat *AP2* gene family (comprising at least 101 members) is consistent with their results, with the additional members falling into a subfamily whose members are most related to APETALA2 (which contains two AP2 domains) and a small subfamily whose members are related to RAV1. Preliminary analysis suggests that the roles of A-subgroup TFs are predominantly involved in the regulation of abiotic stress responses (C-repeat binding factor [CBF] and DREB genes are examples); all of the *AP2* genes that are involved in disease resistance

responses are found in the B subgroup. These findings are based on data from both overexpression experiments and expression patterns. We propose that 12 distinct subfamilies best accommodate the structural diversity within the A and B subgroup, as indicated in Figure 1. These 12 subfamilies were first identified in the *Arabidopsis* gene set. We then mined the indica rice preliminary genome sequence and the wheat unigene set for members of the A and B subgroup (http://ricetfdb.bio.uni-potsdam.de/v2.1/) and confirmed the existence of the 12 subfamilies in both rice and wheat. This suggests that the primary expansion of the A and B subfamily had already occurred when the last common ancestor of monocots and dicots emerged.

AP2 like genes in wheat

AP2-like genes are characterized by having two plant-specific DNA binding motifs referred to as AP2 domains and have been implicated in a wide range of plant development roles. In *Arabidopsis*, *AP2* is a floral homeotic gene involved in the establishment of floral meristem identity (Bowman et al 1993), floral organ identity (Jofuku et al 1994) and temporal and spatial regulation of floral homeotic gene expression (Drews et al 1991). According to our condition for the cDNA library preparation we were identified only three members of this family so it is difficult to draw a good view of this subfamily. In a previous study, it was shown that AP2 subfamily consists of two clades, each of which corresponds to AP2 and ANT groups (Shigyo et al 2006). Genes of the ANT group have the distinctive 10 amino acid residues in AP2 repeat-1 domain (Shigyo et al 2006). The functions of only a few genes of the AP2 and ANT groups have been characterized. Therefore, it is hard to speculate on the evolution of the functions of this gene family. When we consider that previously characterized *AP2* and *ANT* genes are involved in several developmental processes and that the number of these groups increased during land-plant evolution, the evolution of this gene family via gene duplication and subsequent functional diversification is likely to be related to the evolution of developmental processes in

land plants, similar to other transcription factors (Carroll et al 2001). It is interesting to note that *AP2* gene expression in *Arabidopsis* is regulated at the level of translation by a microRNA, which binds to an AASSGF box (Chen, 2004). The AASSGF box is conserved between AP2 in dicots and monocots and was found in one TaAPETALA2.

DREB like genes in wheat

We propose that 6 distinct subfamilies best accommodate the structural diversity within the A subgroup (Figure 1). These six subfamilies were first identified in the *Arabidopsis* gene set. We then mined the indica rice preliminary genome sequence and the wheat unigene set for members of the A subgroup, and confirmed the existence of the 6 subfamilies in both rice and wheat. This suggests that the primary expansion of the A subfamily had already occurred when the last common ancestor of monocots and dicots emerged. Previous analysis suggested that the roles of A-subgroup TFs are predominantly in the regulation of abiotic stress responses (C-repeat binding factor [CBF] and DREB genes are examples). This result shows that CBF group was expanded in wheat after the divergence from dicots. Members within this group may have recent common evolutionary origins and may possess specific functions. Since wheat is a cultivated species, selection either during domestication from its wild ancestor or during agricultural improvement in the subsequent time may also have contributed to the evolution of wheat DREB family.

ERF like genes in wheat

Sakuma *et al.* (2002) divided ERF (B subfamily) into 6 further subfamilies on the basis of conserved regions. Our own analysis of the full *Arabidopsis ERF*gene family, rice and wheat is consistent with their finding analyis. In another study, Gutterson and Reuber (2004) proposed that 10 distinct subfamilies best accommodate the structural diversity within the B subgroup. These 10 subfamilies existed in *Arabidopsis,* rice and tomato. This suggests that the primary expansion of the B

subfamily had already occurred when the last common ancestor of monocots and dicots emerged. Further inspection of the phylogenies, based on more recent amplification events, suggests that as few as 20 B-subfamily genes may have been present in the last common ancestor. ERF genes show a variety of stress-regulated expression patterns. Regulation by disease-related stimuli, such as ethylene (ET), jasmonic acid (JA), salicylic acid (SA) and infection with virulent or avirulent pathogens, has been shown in several ERF genes. However, some ERF genes are also induced by wounding and abiotic stresses. At present, it is difficult to assess the overall picture of ERF regulation in relation to phylogeny because different studies have concentrated on different ERF genes, treatments and time points. Some ERF subgroups are enriched in such genes, suggesting that they have conserved functions that are required for the regulation of disease resistance pathways. Despite broad conservation of their function in monocots and dicots, some structural elements are specialized within each of these two lineages.

Conclusions and directions

Based on results shown in Figure 1 and Table 1, all AP2 subfamilies and subgroups have representatives in monocots. The presence of these groups in *Arabidopsis*, rice, and wheat suggest that the appearance of many of the genes in these species predates monocot/eudicot divergence. Phylogenetic analysis (top portion of Figure 1) reveals a number of monocot AP2 members that have diverged significantly from *Arabidopsis* members. The phylogeny for the DREB/CBF subgroup A-1 (Figure 2) reveals that this group has amplified in wheat to at least 14 classes of genes versus 6 for rice and *Arabidopsis*, suggesting that such events could be the basis for the high freezing tolerance in wheat. The determination of the sequence and chromosome location of the majority of the members of the *CBF* gene family in hexaploid wheat provides the foundation to study the contribution of the

individual *CBF* genes to the observed differential frost tolerance phenotypes. Based on these sequences it is now possible to design *CBF*-specific primers to determine the differential expression of each of these genes under different environmental conditions and study the transcription profiles in freezing tolerant and freezing susceptible lines (Badawi et al 2007). These gene-specific primers can be used to characterize the allelic variation at each of these genes to study their association to varying levels of cold stress in different regions of the world. Phylogenetic and comparative analyses of *AP2* genes in *Arabidopsis*, rice and wheat will act as a first step toward a comprehensive functional characterization of the *AP2* gene family by reverse genetic approaches in the future. The results from the comparative study between *Arabidopsis*, rice and wheat will also provide useful information regarding the functions of *AP2* genes in agronomic, economic, and ecological traits in wheat and possibly in other beneficial plant species. In conclusion, data reported here reveal that multiple AP2 genes are functional in wheat and some members may evolved to play important roles in mechanisms that govern wheat development and adaptability to environmental stresses.

Tables

I-Table 1: AP2 FAMILY IN RICE, *ARABIDOPSIS* AND WHEAT

SUBFAMILY	SUBGROUP	RICE	ARABIDOPSIS [1]	WHEAT	ROLE
APETALA2		20	18	3	Development
RAV		4	6	3	Abiotic stress
CBF/DREB	A - 1	10	6	37	Abiotic stress
	A - 2	6	8	4	
	A - 3	1	1	1	
	A - 4	6	16	2	
	A - 5	24	15	7	
	A - 6	9	10	2	
ERF	B - 1	13	15	23	Biotic stress
	B - 2	15	5	10	
	B - 3	19	18	10	
	B - 4	12	7	1	
	B - 5	6	8	2	
	B - 6	18	12	2	
OTHERS		2	2		
TOTAL		165	147	107	

[1] Classification of *Arabidopsis* members is according to Sakuma et al (2002). Rice and wheat proteins were classified using the highest BLASTP score against the *Arabidopsis* data set.

I-Table 2: List of AP2 cDNAs isolated from hexaploid wheat (cv Norstar)

Gene ID	cDNA (bp)	Protein (a.a)	GI Number	Comments	Reference
TaCBFIVa-A2	950	225	EF028769.1	A-1	1
TaCBFIVa-2.2	822	230	EF028770.1	A-1	1
TaCBFIVa-2.3	660	235	EF028771.1	A-1	1
TaCBFIIIc-3.1	896	227	EF028758.1	A-1	1
TaCBFIIIc-3.2	943	246	EF028759.1	A-1	1
TaCBFIIIc-D3	963	245	EF028760.1	A-1	1
TaCBFIVd-4.1	872	222	EF028781.1	A-1	1
TaCBFIVd-B4	865	222	EF028781.1	A-1	1
TaCBFII-5.1	1069	225	EF028752.1	A-1	1
TaCBFII-5.2	991	219	EF028753.1	A-1	1
TaCBFII-5.3	876	228	EF028754.1	A-1	1
TaCBFIIIa-6.1	958	236	EF028755.1	A-1	1
TaCBFIIIa-6.2	947	242	EF028756.1	A-1	1
TaCBFIIIa-D6	959	238	EF028757.1	A-1	1
TaCBFIVd-9.1	1080	269	EF028782.1	A-1	1
TaCBFIVd-B9	999	269	EF028783.1	A-1	1
TaCBFIVd-D9	1063	269	EF028784.1	A-1	1
TaCBFIIIc-B10	1000	240	EF028761.1	A-1	1
TaCBFIa-A11	962	218	EF028751.1	A-1	1
TaCBFIIId-12.1	1033	245	EF028762.1	A-1	1
TaCBFIIId-B12	970	245	EF028763.1	A-1	1
TaCBFIVc-14.1	1003	212	EF028777.1	A-1	1
TaCBFIVc-B14	863	214	EF028778.1	A-1	1
TaCBFIIId-14.3	892	214	EF028779.1	A-1	1
TaCBFIIId-A15	1010	239	EF028764.1	A-1	1
TaCBFIIId-15.2	926	241	EF028765.1	A-1	1
TaCBFIIId-A19	984	234	EF028766.1	A-1	1
TaCBFIIId-B19	954	234	EF028767.1	A-1	1

Table 2 (continued)

TaCBFIIId-D19	965	234	EF028768.1	A-1	1
TaCBFIVb-A20	1077	217	EF028772.1	A-1	1
TaCBFIVb-B20	887	212	EF028773.1	A-1	1
TaCBFIVb-D20	980	212	EF028774.1	A-1	1
TaCBFIVb-21.1	979	202	EF028775.1	A-1	1
TaCBFIVb-D21	1066	202	EF028776.1	A-1	1
TaCBFIVd-A22	1219	275	EF028785.1	A-1	1
TaCBFIVd-B22	1252	290	EF028786.1	A-1	1
TaCBFIVd-D22	1211	275	EF028787.1	A-1	1
TaBBM2	1746	375	*	Apetala2	2
TaAPETALA2	1730	355	*	Apetala2	2
TaANTL1B	1577	387	*	Apetala2	2
TaRAV1	1524	382	*	RAV1	2
TaRAV1	1373	392	*	RAV1	2
TaRAV2	1004	264	*	RAV2	2
TaDREB3B	1432	264	*	A-2	2
TaDREB4B	1699	364	*	A-2	2
TaDREBP2	1289	278	*	A-2	2
TaDREBP2	1275	254	*	A-2	2
TaDREBP4A	972	232	*	A-3	2
TaCBF7	964	275	*	A-4	2
TaCBF7	902	277	*	A-4	2
TaTINY	1220	350	*	A-4	2
TaDREBP	1275	243	*	A-5	2
TaDREBP	958	230	*	A-5	2
TaDREBP	967	232	*	A-5	2
TaDREBP	1342	239	*	A-5	2
TaDREBP3	1882	351	*	A-5	2
TaDREBP3	1720	328	*	A-5	2
TaDREBP	874	179	*	A-5	2
TaDREB3	1845	337	*	A-6	2
TaDREB2	1679	284	*	A-6	2
TaERF2-1	1016	271	*	B-1	2

Table 2 (continued)

TaERF2-2	1074	272	*	B-1	2
TaERF2-3	1268	236	*	B-1	2
TaERF2-4	1166	198	*	B-1	2
TaERF2-5	922	193	*	B-1	2
TaERF3-1	902	234	*	B-1	2
TaERF3-2	984	234	*	B-1	2
TaERF3-3	817	206	*	B-1	2
TaERF3-4	767	219	*	B-1	2
TaERF2-6	979	236	*	B-1	2
TaERF2-7	960	240	*	B-1	2
TaERF3-5	1055	248	*	B-1	2
TaERF7-1	977	193	*	B-1	2
TaERF7-2	824	215	*	B-1	2
TaERF7-3	958	195	*	B-1	2
TaERF7-4	1180	210	*	B-1	2
TaERF-1	1061	320	*	B-1	2
TaERF-2	1174	316	*	B-1	2
TaERF-3	1283	316	*	B-1	2
TaERF-4	1370	316	*	B-1	2
TaERF-5	1201	313	*	B-1	2
TaERF3-5	698	219	*	B-1	2
TaEREBP	1108	222	*	B-1	2
TaERF1-1	1335	308	*	B-3	2
TaERF1-2	1029	279	*	B-3	2
TaERF1-3	1455	377	*	B-3	2
TaERF1-4	1545	383	*	B-3	2
TaERF1-5	1805	346	*	B-3	2
TaERF1-6	1483	308	*	B-3	2
TaERF1-7	1486	348	*	B-3	2
TaERF1-8	742	185	*	B-3	2
TaERF1-9	879	184	*	B-3	2
TaERF2-1	1004	244	*	B-3	2
TaEREBP2-1	1061	285	*	B-3	2

Table 2 (continued)

TaEREBP2-2	1200	239	*	B-3	2
TaDREBP1-1	1588	398	*	B-2	2
TaEREBP1-2	1469	355	*	B-2	2
TaEREBP1-3	1486	355	*	B-2	2
TaEREBP1-4	1403	355	*	B-2	2
TaEREBP1-5	1587	381	*	B-2	2
TaEREBP1-6	1523	380	*	B-2	2
TaEREBP1-7	1758	383	*	B-2	2
TaEREBP1-8	1501	398	*	B-2	2
TaEREBP1-9	1222	303	*	B-2	2
TaEREBP1-10	1381	306	*	B-2	2
TaEREBP1-11	1438	342	*	B-2	2
TaEREBP1-12	984	356	*	B-2	2
TaERF4-1	1327	336	*	B-1	2
TaRAP2.6	1011	286	*	B-4	2
TaERF4-2	1227	336	*	B-1	2
TaERF4-3	1019	208	*	B-1	2
TaWR	1117	265	*	B-5	2
TaSHINE1-1	786	222	*	B-6	2
TaSHINE1-2	1087	227	*	B-6	2

1= Badawi et al. 2007

2= this study

ORF = open reading frame (CDs)

bp = base pair

a.a = amino acid

* = sequences are presented in Supplemental data and will be submitted to GenBank.

I-Table 3: Mapping of 20 wheat AP2 EST and genes

Gene ID	Chromosomal location	
	Arm	Bin
TaAPETALA2	5AL	5AL17-0.78-0.87
	5DL	5DL5-0.76-1.00
TaANTL1B	7AL	C-7AL1-0.39
	7BL	C-7BL2-0.33
	7DL	C-7DL5-0.30
TaANTL	4AS	4AS3-0.76-1.00
	4BL	C-4BL5-0.71
	4DL	4DL12-0.71-1.00
TaCBFIVa-A2 (A-1)	5AL	5AL10-0.57-0.78
	5BL	5BL1-0.55-0.75
	5DL	5DL1-0.60-0.74
TaCBFIIIc-D3 (A-1)	5AL	5AL10-0.57-0.78
	5BL	5BL1-0.55-0.75
	5DL	5DL1-0.60-0.74
TaCBFIVd-B4 (A-1)	5AL	5AL12-0.35-0.57
	5BL	5BL1-0.55-0.75
	5DL	5DL1-0.60-0.74
TaCBFIIIa-6.1 (A-1)	5AL	5AL10-0.57-0.78
	5BL	5BL1-0.55-0.75
	5DL	5DL1-0.60-0.74
TaCBFIVc-B14 (A-1)	5AL	5AL12-0.35-0.57
	5BL	5BL1-0.55-0.75
	5DL	5DL1-0.60-0.74

Table 3 (continued)

Gene ID	Chromosomal location	
	Arm	Bin
TaCBFIIId-15.2(A-1)	5AL	5AL10-0.57-0.78
	5BL	5BL1-0.55-0.75
	5DL	5DL1-0.60-0.74
TaCBFIVb-D21 (A-1)	5BL	5BL1-0.55-0.75
TaCBFIVd-A22 (A-1)	5AL	5AL10-0.57-0.78
	5BL	5BL1-0.55-0.75
	5DL	5DL1-0.60-0.74
TaDREB3B (A-2)	1AL	1AL1-0.17-0.61
	1BL	C-1BL6-0.32
	1DL	1DL2-0.41-1.00
	2BL	2BL6-0.89-1.00
TaDREB3(A-6)	6AL	6AL4-0.55-0.90
	6BL	6BL5-0.40-1.00
	6DL	6DL1-0.47-0.68
TaERF7-2(B-1)	2A	C-2AL1-0.85
	2B	2BL4-0.50-0.89
TaEREBP1-3 (B-2)	5BL	5BL1-0.55-0.75
	5DL	C-5DL1-0.60
TaEREBP1-4 (B-2)	5BL	C-5BL14-0.75
	5DL	C-5DL1-0.60
TaERF1-2 (B-3)	4DS	C4DS1-0.53

Table 3 (continued)

Gene ID	Chromosomal location	
	Arm	Bin
TaERF1-4 (B-3)	7AS	7AS5-0.59-0.89
	7BS	7BS1-0.27-1.00
TaERF1-7 (B-3)	3AS	3AS4-0.45-1.00
	3BS	3BS9-0.57-0.78
	3DS	3DS6-0.55-1.00
	1AS	1AS1-0.47-0.86
TaRAP2.6L (B-4)	2AL	C-2AL1-0.85
	2BL	2BL2-0.36-0.50
	2DL	C-2DL3-0.49

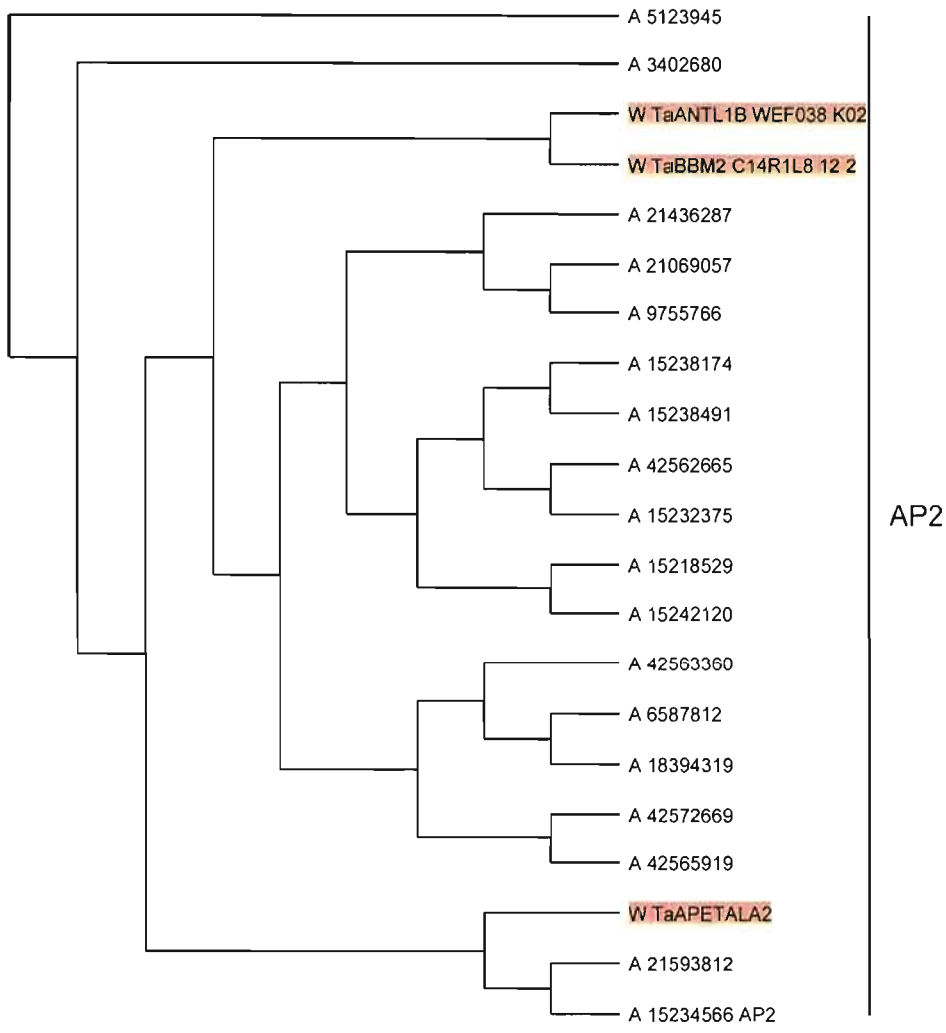
Figures

I- Figure 1: Phylogenetic tree of AP2 genes from wheat and *Arabidopsis*. Deduced amino acid sequences of the AP2 domain from 147 *Arabidopsis* genes and 78 wheat genes were aligned with ClustalW, and the neighbor-joining trees were produced using PAUP 4.0 (Swofford 2003). Since wheat is hexaploid, only one copy of homeologous proteins was included. Members of each subfamily, DREB, ERF, APETALA2 and RAV, were classified with a few representative sequences from other subgroups.

Colored rectangles indicate the wheat proteins. On the right side is indicated the subfamily and subgroup of the corresponding *Arabidopsis* AP2 proteins.

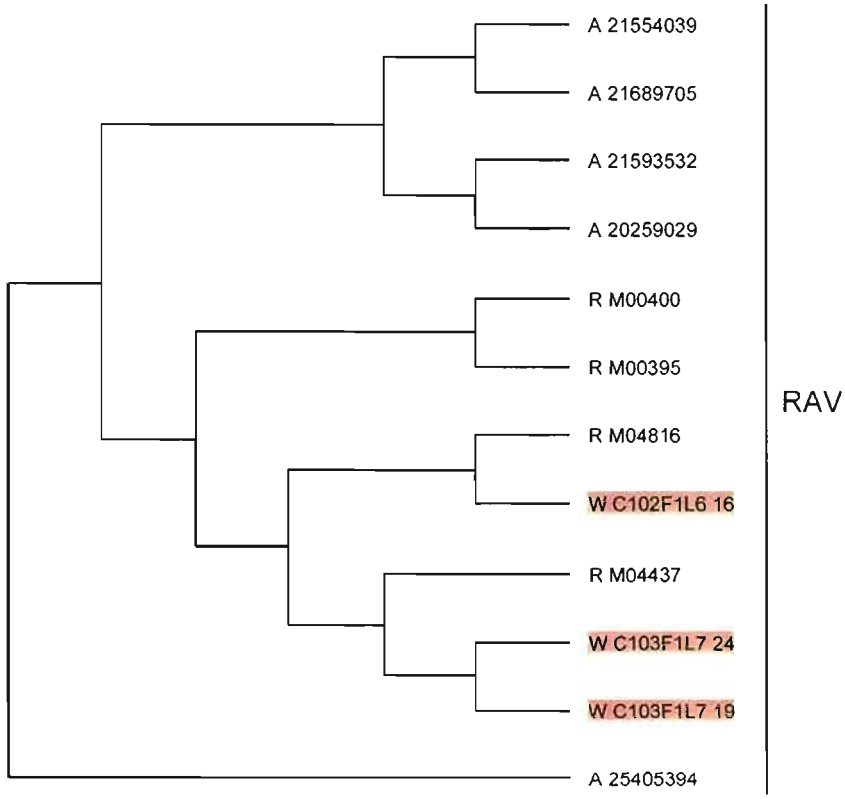
A *Arabidopsis thaliana*, W *Triticum aestivum*

.


A 5123945
A 3402680
W TaANTL1B WEF038 K02
W TaBBM2 C14R1L8 12 2
A 21436287
A 21069057
A 9755766
A 15238174
A 15238491
A 42562665
A 15232375
A 15218529
A 15242120
A 42563360
A 6587812
A 18394319
A 42572669
A 42565919
W TaAPETALA2
A 21593812
A 15234566 AP2
AP2

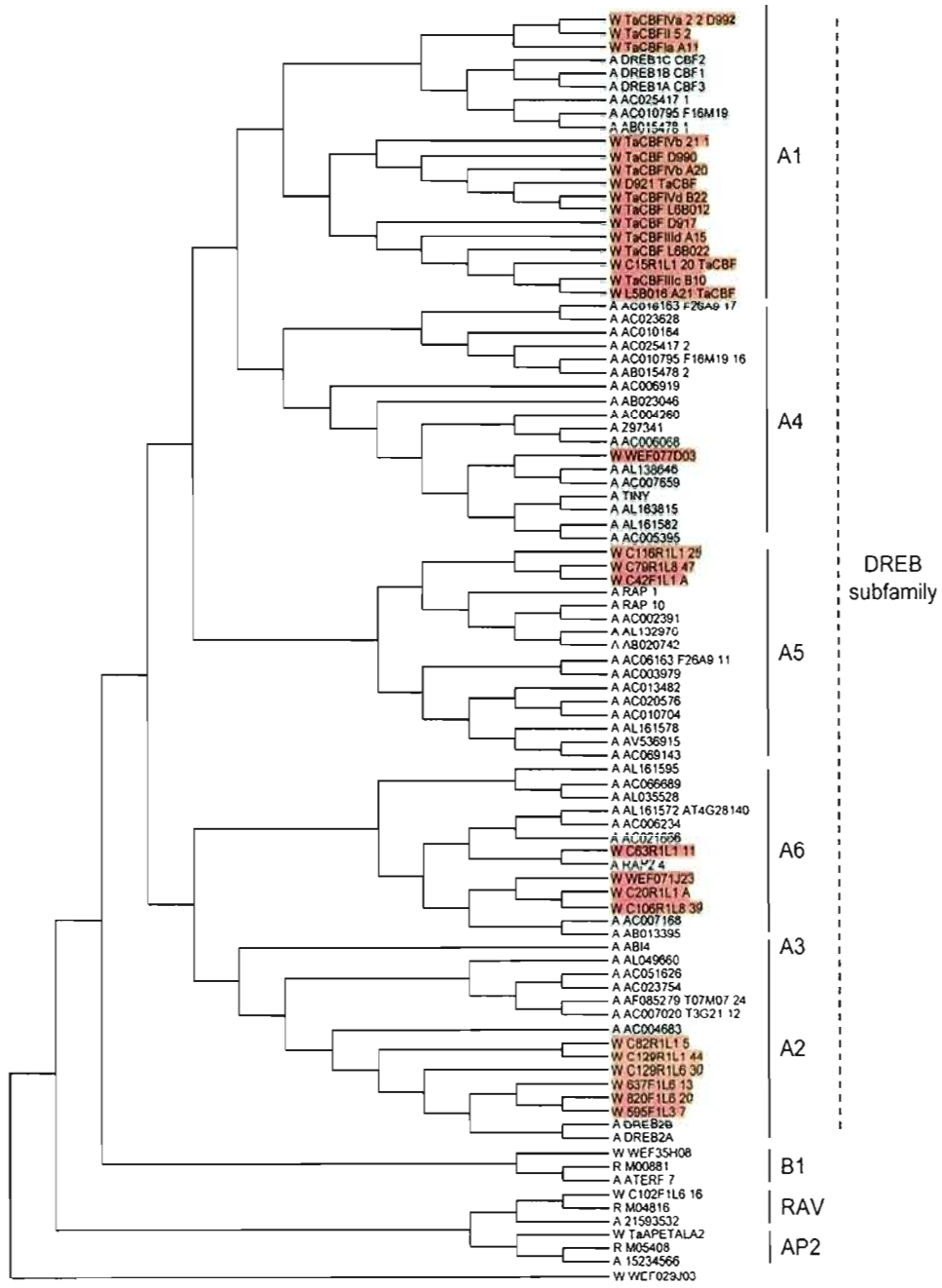
I- Figure 2: Phylogeny of the RAV subfamily. Colored rectangles indicate the wheat proteins. On the right side is indicated the subfamily and subgroup of the corresponding *Arabidopsis* RAV proteins.

A *Arabidopsis thaliana*, W *Triticum aestivum*, R *Oryza sativa*.


A 21554039
A 21689705
A 21593532
A 20259029
R M00400
R M00395
R M04816
W C102F1L6 16
R M04437
W C103F1L7 24
W C103F1L7 19
A 25405394
RAV

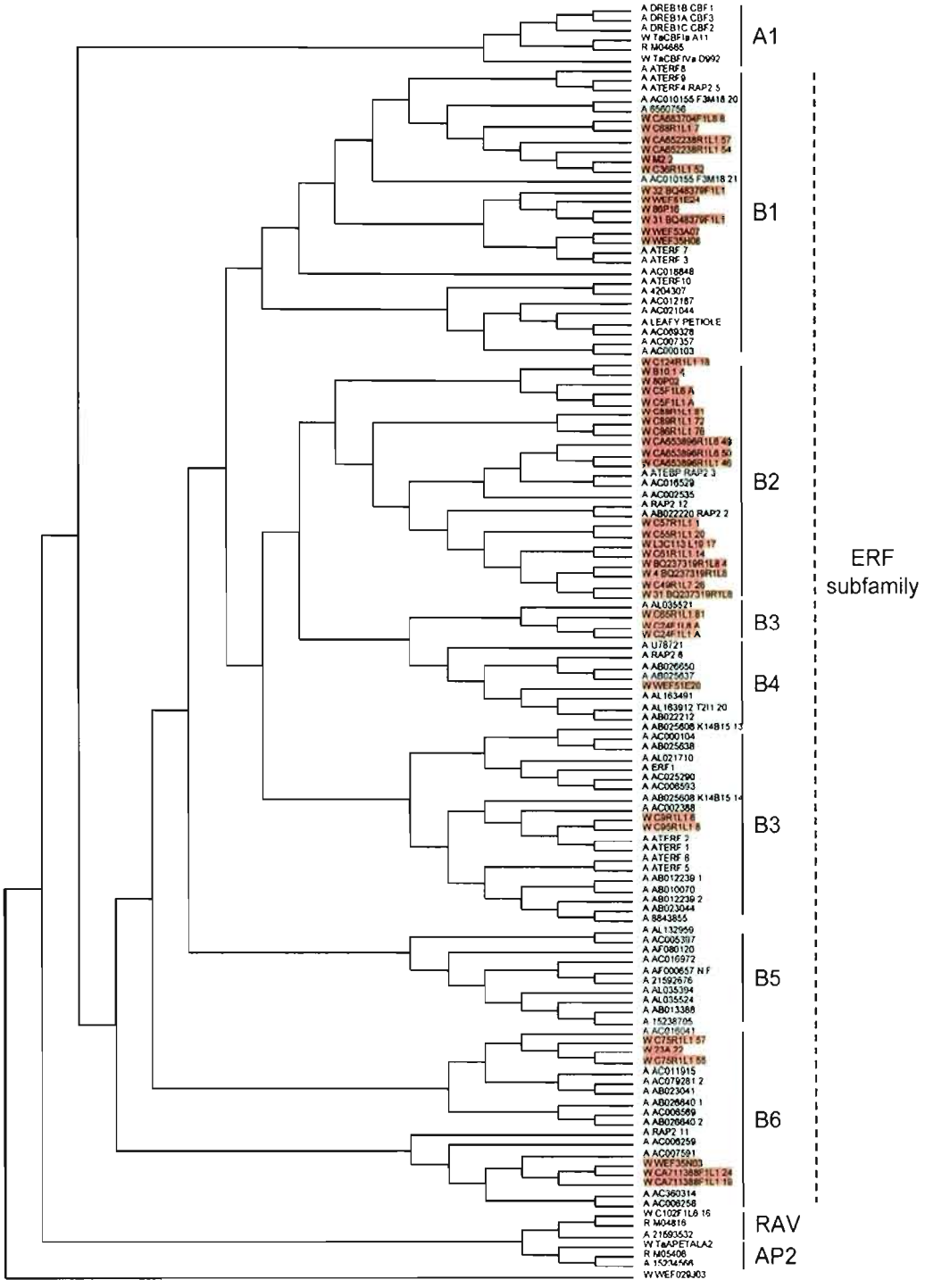
I- Figure 3: Phylogeny of the DREB subfamily. Colored rectangles indicate the wheat proteins. On the right side is indicated the subfamily and subgroup of the corresponding *Arabidopsis* DREB proteins. Several genes from rice were also involved as representatives. WEF029J03 from wheat was used as outgroup.

A *Arabidopsis thaliana*, W *Triticum aestivum*, R *Oryza sativa.*


W TaCBFIVa 2 2 D992
W TaCBFII 5 2
W TaCBFIa A11
A DREB1C CBF2
A DREB1B CBF1
A DREB1A CBF3
A AC025417 1
A AC010795 F16M19
A AB015478 1
W TaCBFIVb 21 1
W TaCBF D990
W TaCBFIVb A20
W D921 TaCBF
W TaCBFIVd B22
W TaCBF L8B012
W TaCBF D917
W TaCBFIIId A15
W TaCBF L8B022
W C15R1L1 20 TaCBF
W TaCBFIIIc B10
W L5B016 A21 TaCBF
A AC016163 F26A9 17
A AC023628
A AC010164
A AC025417 2
A AC010795 F16M19 16
A AB015478 2
A AC006919
A AB023046
A AC004260
A Z97341
A AC006068
W WEF077D03
A AL139646
A AC007659
A TINY
A AL163815
A AL161582
A AC005395
W C116R1L1 25
W C70R1L8 47
W C42F1L1 A
A RAP 1
A RAP 10
A AC002391
A AL132970
A AB020742
A AC06163 F26A9 11
A AC003979
A AC013482
A AC020576
A AC010704
A AL161578
A AV536915
A AC069143
A AL161595
A AC066689
A AL035528
A AL161572 AT4G28140
A AC006234
A AC021696
W C83R1L1 11
A RAP2 4
W WEF071J23
W C20R1L1 A
W C106R1L8 39
A AC007168
A AB013395
A ABI4
A AL049660
A AC051626
A AC023754
A AF085279 T07M07 24
A AC007020 T3G21 12
A AC004683
W C82R1L1 5
W C129R1L1 44
W C129R1L6 30
W 837F1L8 13
W 820F1L8 20
W 595F1L3 7
A DREB2B
A DREB2A
W WEF35H08
R M00881
A ATERF 7
W C102F1L6 16
R M04816
A 21593532
W TaAPETALA2
R M05408
A 15234566
W WCF023J03
A1
A4
A5
A6
A3
A2
B1
RAV
AP2
DREB subfamily

I- Figure 4: Phylogeny of the ERF subfamily. Colored rectangles indicate the wheat proteins. On the right side is indicated the subfamily and subgroup of the corresponding *Arabidopsis* ERF proteins. Several genes from rice were also involved as representatives. WEF029J03 from wheat was used as outgroup.

A *Arabidopsis thaliana*, W *Triticum aestivum*, R *Oryza sativa*.


A1
B1
B2
B3
B4
B3
B5
B6
RAV
AP2
ERF
subfamily

I-Supplemental DATA

I-Table S1: List of AP2 cDNAs sequences isolated from wheat cv Norstar (*Ta APETALA2* to *TaEREF*).

>**TaBBM2 clone** C14R1L8-12, 1746bp CDS complete, mRNA 375-**1502**
CGTTTCCTCTCTCGATTTGTCCAAATCTTTTGTTCCTCCTCCTCCACCAGCGATTAGTTTGTTGTTTCCG
GCATCACTCCGCACTAGGCCGCCCGCGCTGGCCTCGTCGTATCCTTCCCCAATTCCGCCGCCCCTCCGCG
CCCGATATTTATTTCCTGCCTCAGCATCCATTTCCGTTGATAGATTTTTCCAGCTTTTGCTGCCTCGCCG
TCGCTGCTAATATCCGCGCTGGGATATTTCTTCTTTTGCTCCGGAGGCCTCCGGATCTTTCGATCGCGGC
GAGCGGTCGGCTCAAGGTAGTTCGTGAATAGGAAGGCTGATAGGCAGGTTATTAGGGTTTTGGGAGTTGT
TTTTGTCTCTGCTCCAATACATAG**ATG**ATGAAATCCGGGGAGGAAGTTAGTCAGGGTCAGCAAATGAACG
ATTTTTCGGAGGAGAAAGCTGCTGGGGAATCGGGGGATGGTCGGAAGATCGAGAGGAGCCCTTCTATCAA
TCTGAATTCCTTGCCTGCAATAGCCCCTGCCACTACGGAGATTGGTGTCTTGCACGGCGCAGTGGAGTCA
GAGGCCAACGATGCAAGCACTCACAAGGGAGACGAGTCCAGTGGCACTGATCAGAAGAAGGTTCTGAAGA
ATGAAGAAGTTGATGAAGCTGAAGTTCAGGCCTGTGCAGATGTGAAGAGCGACTCGGTTGACCCTTTGAA
CAGCGAGAACCATGCCGGGGAGAAGGATGCTTTGGTAACTGTGGCAGAAAATGAGGGGTGTGCGGATGGT
GGCGATAATTATAAGGGAGTTCAAGTTCTCAGCATTGTCAAAAAGGATGAGTCTGAGGAAATTGTTGATT
CTATTAATCCTGTGACAGTTGCGGGGTATAGAGAGGAGAAGGGCGCCTCTGGTTCTACTTCTGCAGTTAC
TGCGGTGCGAGCACCTGGCTCCCGCTCATCTTGTTTCCATGGTGTGACCAGGCATAGGTGGAGTGGGAAA
TATGAAGCTCATTTGTGGGACAGTACTTGCAGAGTAGAAGGACGGAGAAGGAAAGGGAAGCAAGTTTATT
TAAGAAGTTATGATACTGAGCAAAAAGCTGCCAGGGCATATGATGTTGCAGCTCTTAAATTCTTTGGACT
AAATACAAAGCTGAACTTTTCGATTTCGGAATATGAGAAGGAACTGGCGGACATACAAGACATGTCTCCA
GAGGAATGTGTGACATACTTGAGAAGGAGGAGTAGCTGCTTCTCAAGAGGGGCGTCTATTTACAGAGGAG
TTACAAGGAGGCAGAAAGATGGTCGATGGCAGGCACGCATAGGACTGATTGCTGGAACTAGAGACATTTA
CCTTGGAACTTTCAAAACTGAGGAAGAAGCCGCAGAAGCTTATGATATTGCTGCCATCGAGATACGCGGC
AAAAACGCGGTGACCAACTTTGACAGAAGCAACTACATGGACAGGGGCATGCATTGTATAGAAGGCGCAG
GGTTGAAGCTGCTTGCAACCAAGCCAGAA**TAG**TACCTGATTTGGTATCGTATATTGAACAGATTTGGTTG
GCCGTATTTTGGAGCCTAGTGGTACATACAGATAGAAGAACTGGTTGCAGCCTGTCATTATCTGCTGCTG
TATGATTCGTCAGATTATATATAGTTCTTTCAGAGAGAATTTCAGTAATTTAGCATGCTTTGTGTCCAGA
ACAAGATTTTGACCATGCATTACTGTTATAGTGTTTGTAGGCTAGAGTTGCAGTGGAAGATGTTGC

>**TaAPETALA2 clone** WEF048_C12, 1730bp CDS complete, mRNA **129-1472**

GCGGCCACCGCGCTCCCATGCCATAGACGCGACCCCACTCATCGGTCCAGGTCGGTCGCTCGGAGCCGAG
CGGCGGCGGGCGGGCGAGGAGTGCGTTTTATTCGGTCCCGGCGGGCCTCGGATCGGAG**ATG**GTGCTGGAT
CTCAATGTGGAGTCGCCGGCGGACTCGGGCACGTCCAGCTCCTCCGTGCTCAACTCCGCGGACGCCGGTG
GCGGCGGCTTCCGGTTCGGCCTGCTCGGGAGCCCTGATGATGACGACTGCTCCGGCGAGCCGGCGCCGGT
CGGGCCCGGGTTCGTCACGAGGCAGCTCTTCCCCGCGTCGCCGCCCGGGCACGCGGGCGCGCCCGGGGTG
ACGATGGGGCAGCAGGCCCCGGCGCCTGCGCCGATGGCGCCCGTGTGGCAGCCGCGGCGCGCCGAGGAGC
TCCTCGTGGCGCAGCGGATGGCGCCCGCGAAGAAGACGCGGCGGGGCCCGAGGTCGCGCAGCTCGCAGTA
CAGGGGCGTCACCTTCTACCGCAGGACCGGCCGGTGGGAGTCGCACATCTGGGATTGCGGGAAGCAGGTC
TACTTGGGTGGTTTCGACACTGCGCACGCGGCCGCAAGGGCCTACGATCGCGCGGCGATCAAGTTCCGGG
GGCTGGAGGCCGACATCAACTTCAATCTGAGCGACTACGAGGAGGATTTGAAGCAGATGAGGAACTGGAC
CAAGGAGGAGTTCGTGCACATCCTCCGCCGCCAGAGCACGGGGTTCGCCAGGGGGAGCTCCAAGTACCGC
GGCGTCACGCTCCACAAGTGCGGCCGCTGGGAGGCAAGGATGGGCCAGCTGCTCGGCAAGAAGTACATAT
ATCTGGGCCTCTTTGACAGCGAAGTTGAAGCTGCAAGGGCGTACGACAGGGCGGCGATTCGCTTCAATGG
GAGGGAAGCTGTGACTAACTTTGAGAGCAGCTCCTACAATGGGGATGCTCCACCCGACGCCGAAAATGAG
GCAATTGTTGATGCTGATGCTCTTGACTTGGATCTGCGGATGTCGCAACCCACCGCGCACGATCCCAAGA
GGGACAACATCATCGCCGGCCTTCAGTTAACTTTTGATTCCCCTGAATCGTCAACCACAATGATCTCTTC
TCAGCCAATGAGCTCATCTTCGTCCCAGTGGCCTGTGCATCAACATGGCACGGCAGTAGCACCTCAGCAG
CACCAGCGTTTGTACCCATCTGCTTGTCATGGCTTCTACCCGAACGTACAGGTGCAGGTGCAGGAGAGGC
CCATGGAGGCAAGGCCCCCTGAGCAGCCGTCGTCCTTCCCCGGCTGGGGGTGGCAAGCGCAAGCCATGCC
GCCGGGCTCCTCCCACTCGCCGTTGCTTTACGCTGCAGCATCATCAGGATTTTCTACCGCCGCCGCCGGC

GCGAACCTCGCCCCGCCGCCGCCGTACCCGGACCACCACCGGTTCTACTTCCCCCGCCCGCCGGACAACT
GAAGCTGGCCGTTGTGACCAGACGGCGGTGGGTGCGCGCGGTCGAGGTGTTCGCTCCTCGTCGTCGGTAA
CGCTTGTTGTGAAACTATAATCGGAGAGAGATGACATTGCCAGGCCATGTGTGGTGACACTACTGGCTGG
TCTCTCGCCGCCTCGCCATGATCGGGATCACGCGGATCATGGCTGTTCATTAGATTCTCATGTATCCAAT
GTTCAAGTTTCCCAAACGGTTGAAAAAACTTTGAAATTTGTGATGGCAAA

>**TaANTL1B clone** WEF038_K02, 1577bp CDS complete, mRNA **114-1277**

CGACGCAGTCACGCACCGGCTCTTTCCTCCACCCCTCGTAGGGCTAGCTAGGGTTTCTCCGCCCAAATCG
TGTAATAATAACCCACTTGAAGATTCGCGCGAACGCTGCGGCGATGGCCACCACCGTCCAACCCCACTCC
CCGGACCCAACAGCCATCACCACCACCCCTGCCCCACCCCCTCCCCCTTCTCCACCTCGCCAGGAGAACC
CGACCGCCGCGGGGGAAGGCGTAGAGATCGCGGCGCTCGATGAGCAGCCTGCCGCCGTCGCCGTCGCCGA
CAAGGGGAAGACGGCCCCCGGCGGCGGGAAGCTGGTGGCGGAGGCCATGCGCAAGTGCGCGGCGCCCCGG
TCGTCGCGCTACCACGGCGTGACGAGGCTCAAGTGGAGCGGCAAGTACGAGGCACACCTCTGGGACAACA
CCAGCCAGGTTGAGGGGCGCAAGCGCAAGGGCAAGCATGTGTACTTGGGAAGCTATGTTACTGAAGAGAA
TGCTGCAAGGGCACATGACCTTGCAGCCCTGAAATATTGGGGCATAACTCAACCCACCAAACTAAACTTC
AATATTTCTGATTATGCAAAAGAAATTGAGATCATGAAGAGCATGAATCAAGATGAATTTGTGGCCTACA
TAAGGAGGCAGAGTAGTTGTTTCTCAAGAGGAACATCGTCATACAGGGGTGTAACAAGACGAAAGGATGG
TAAATGGCAAGCACGTATTGGTAGGATTGGTGAGAGTAGAGACACTAAAGACATCTATCTTGGGACCTTT
GAAACTGAAGTGGAGGCAGCTGAAGCGTATGACCTAGCAGCAATTCAGCTCCGTGGTGTTCATGCTGTGA
CCAACTTTGATATCAGCAACTACTCCGAAGAAGGTTTGAAGAAACTAGAAGGCTCATCCGAGGTAGTGAA
CCTGGAGGACCAATCAGAAGTCACTAAGTTAGCTGTGACCAACTTGGATATTAGCAAACACTGCGAAGAT
GGTTTGAAGAAACTAGATGGCGCATCCCAGATAGTGAACCTGGAGGACCAATCAGAAGTCACCAAGTTAT
CTGTGACCAACTTTGATATTAGCAACTGCTGTGAAGATGGTTTGAAGAAACTAGATGGCGCATCCCAGAT
AGTGAACCTGGAGGACCAATCAGAAGTCACCAAGTTATCTGTGACTAACTTTGATATTAGCAACTGCTGT
GAAGATGGTTTGAAGAAACTAGAAGGCTCATCCGAGGTAGCGAACCTGGAGGACCAATCAGAAGTCACAA
AGTTAGCTGGACAATAGATATTAGAATAGCAACATGTAAATTATTTATTTCTTCCTTTTTATCTTTTCTC
TGATCGTCAGTCTCTCCACCTTTCTCTTTTCTTGATTGATGATGAGAGTGTCTTTTCTTCGTCCACCTTT
CCTGAAACTTTTTTTTNCCCAAGGAATTTCCCTTCGCTGGTTGCCAACTCTATTATTTACCAGCTCCAGA
GTAGGATCGTGTGATGCTGTACTCTGTATTCTCATGTACACTGATTACAAACCATTGATAGTATTTGTAG
ACATTACATCACAAAGGAAGTTCATTCTTCTATTTTT

>**TaRAV1-1 clone** C102F1L6-16, 1524bp CDS complete, mRNA 20-**1163**

CAGTCGCGCCGTTACTGAAATGGACAGCACCAGCTGCCTCGCCGACGACACCAGCAGCGGCGGCGGCGCC
TCCACGGACAAGCTCAAGGCGCTGGCCGCCGCGGCGGCTGCGGCCGCGGGGCCGCTCGAGCGCATGGGCA
GCGGCGCCAGCGCGGTGCTCGACGCGGCCGAGCCGGGCTCCGAGGCCGACTCTGGCGGCCGTGCGGGTGC
CGTCGGCGCGGCGGCGGGGAAGCTGCCGTCGTCCAGGTTCAAGGGCGTGGTGCCGCAGCCCAACGGGCGG
TGGGGCGCGCAGATCTACGAGCGGCACCAGCGCGTGTGGCTCGGCACGTTCGCCGGGGAGGCCGACGCCG
CGCGCGCCTACGACGTCGCCGCGCAGCGCTTCCGCGGACGCGACGCCGTCACCAACTTCCGCCCGCTCGC
GGACGCCGACCCCGACGCCGCCGCCGAGCTCCGCTTCCTCGCCGCGCGCTCCAAGGCCGAGGTCGTCGAC
ATGCTGCGCAAGCACACCTACTTCGACGAGCTCGCCCAGAGCAAGCGCGCCCTCGCCGCCTCGGCCGCGC
TCTCCGCGCCAACCACATCACGCGGCGCCGCCTCGACCCCCTCGCCGGCCGCCGCGCGCGAGCACCTGTT
CGACAAGACGGTCACGCCCAGCGACGTGGGCAAGCTGAACCGGCTGGTGATTCCGAAGCAGCACGCCGAG
AAGCACTTCCCGCTGCAGCTCCCGGCCGCGGGCGGCGAGAGCAAGGGCCTGCTCCTGAACTTCGAGGACG
CCGCCGGCAAGGTGTGGCGGTTCCGCTACTCGTACCGGAACAGCAGCCAGAGCTACGTCCTCACCAAGGG
CTGGAGCCGCTTCGTGAAGGAGAAGGGCCTTGGCGCCGGAGACGTTGTCGGGTTCTACCGCTCCGCCGCC
GGGAGCACCGGCGAAGACACCAAGCTCTTCATTGACTGCAAGCTGCGACCGGACACCAACAGCCCCGCCT
CCGCTGACCCCGTGGACCAGTCGGCACCTGTGCAGAAGGCCGTGAGACTCTTCGGCGTCGACCTTCTGAC
AGCGCCGGCGTCGCCGGAGCAGGGGATGCCGGGGTGCAAGAGGGCCAGAGACTTGGTGAAGTCGCCGCCT
CCGAAAGTGGCGTTCAAGAAGCAATGCATAGAGCTGGCGCTAGCGTAGAGTTGGAACTATTAGCTCGATC
GATCTCTTCTCTCCAGCTAGGCGGCGGTTTTGCTCGCATAATTCAGGTGGTAGAGCTTAGCTTAATTAGT
CCCTTGTTGGTAGTACCTATCATCAACTTTGTTTGTTTATTTGTCATGTTGTTGCATGCCCTGATGTAAA
TCTTATCCTCCCAAAAATGTATACTAATTCAGGAAGCCCTCGAAGGCTAGCTTTAGATCGTTCAACCGAC
GACATAATAATAACATTTTGAATGTAGCCTAGCTCATCAGTTTCCTTGTCAAAAACTGAAGCATATCCTA
TTAAGTAGTCCAAGTAGTGAGTTATATGATAACCTAGGCAAAAAAAAAAAAAAA

>**TaRAV1-2 clone** C103F1L7-19 1373bp CDS complete, mRNA 61-**1239**

CCCGAGCTAAGCATCTTCTTGATTTCTCGGTGATCTTGATTTCTCCGTGATCGGATTCGG**ATG**GACAGCG
CGAGAAGTTGCCTCGTGGACGACGTGAGCAGCGGCGCGTCCACGGGGAAGAAGGCGTCTCCGGCGCCGGC
AGCGCCGGCGGCCAAGCCGCTGCAGCGCGTGGGCAGCGGGGCCAGCGCGGTCATGGACGCGCCGGAGCCC
GGCGCGGAGGCGGACTCCGGCCGCATCGGCAGGCTGCCGTCCTCCAAGTACAAGGGCGTGGTGCCGCAGC
CCAACGGGCGGTGGGGCGCGCAGATCTACGAGCGCCACCAGCGCGTCTGGCTCGGCACGTTCACGGGGGA
GGCCGAGGCTGCGCGCGCCTACGACGCGGCGGCGCAGCGCTTTCGCGGCCGCGATGCAGTCACCAACTTC
CGCCCGCTCGCCGAGTCCGACCCGGAGGACGCCGCCGAGCTCCGCTTCCTCGCCGGCCGCTCCAAAGCCG
AGGTCGTCGACATGCTGCGCAAGCACACCTACCCCGACGAGCTCGCTCAGTACAAGCGCGCTTACTTCGT
CGCCGCTGCGGCGTCCTCCCCTACATCGTCCTCGTTGCCTCCCGCCTCGTCGCCCTCTTCGGCGGCTGCA
CACTCGCCGGCGGCGCGGCGCGAGCACCTGTTCGACAAGACGGTCACGCCCAGCGACGTGGGGAAGCTGA
ACCGGCTGGTGATACCGAAGCAGCACGCCGAGAAGCACTTCCCTCTCCAGCTTCCTTCCGCCGGCGCCGC
CGTGTCCGGCGAGTGCAAGGGCATGCTGCTCAACTTCGACGACGCGGCCGGCAAGGTGTGGAGGTTCCGG
TACTCGTACTGGAACAGCAGCCAGAGCTACGTGCTCACCAAGGGGTGGAGCCGCTTCGTCAAGGAGAAGG
GCCTGCACGCGGGCGACGCCGTCGGGTTCTACCGCTCCGCCTCCGGCAACAACCAGCTCTTCATCGAGTG
CAAGCTGCGGTCCAAGACCACGACGACGACGACGACCTTCGTCAACGCGGCGGCCGCCGCGTCGCCTGCA
CCGGTGATGAGGACCGTGCGACTCTTTGGCGTCGACCTTCTCACGGCGCCGGCGCCGAGTCACGCGCCCG
AGCACGAGGACTGCAGCATGGTGCCCAAGACATACAAGAGATCCATGGACACCAGCGCAGCGCCCACTCC
GGCGCACGCGGTCTGGAAGAAGCAGTGCATAGACTTCGCGCTGACC**TAG**CTAGCTAGCGTTTTCCCTCCA
TGGTTGCTTTGCTTGCTTCCAAATTTCCATGGCAGTAGCTTAGAGCTCTTGATCGATCCAAGTGTTTGCT
CCTTATATTGAGTTGTTTTCAAACACAACCAAAAAAAAAAAAA

>**TaRAV2 clone** C97R1L1-31, 1004bp CDS complete, mRNA 186-**900**

TACACACAACACCATCTTGCCTCTCCCCTCCTCTCTTCTCTTCTCTCCACTGCCCAACTTTTGCCATAGA
CAAGAACCTCTAGGTAGCTTTCCTCCAGAGCAGCA**ATG**GCCATGCACCCTCTCTCTCAGGGGCACCCACA
GGCCTGGCCATGGGGGGTAGCCATGTACACGAACCTACACTACCACCACCAGTACGAGAGGGAGCACCTG
TTCGAGAAACCCTTGACGCCCAGTGATGTGGGCAAGCTCAACAGGCTGGTGATTCCCAAGCAGCACGCGG
AGAGGTACTTCCCCCTGAACGGCGGCGACTCCCCCGGCGAGAAGGACCTGCTCCTGTCTTTCGAGGATGA
GGCCGGCAAGCCGTGGCGGTTCCGGTACTCGTACTGGACGAGCAGCCAGAGCTACGTGCTCACCAAGGGC
TGGAGCAGGTACGTCAAGGAGAAGCAGCTCGACGCCGGAGACATCGTGTACTTCGAGAGGGTGCGCGGCC
TTGGCACAGGCGACCGGCTCTTCATCGGCTGCAGGCGTCGTGGCGAGAGCGCGCCGCCACCACCTGTACG
TGTACCTCCGCCCGCTCTGAACGCCGGGGAGCAACAGCCTTGGAGCCCGATGTGTTACAGCACGTCAGGA
TCATACCCTACCAGCCCTGCCAATTCCTACGCCTATCGCCGCTCGGTGGAGCAAGATCACAGCGACATGC
TGCATGCAGGCGATTCGCAGAGAGAGGCAGACGCCAAGAGCAGCACGGCATCGGCGGCGCCGTCGAGACG
TCTCCGGCTGTTCGGCGTCAACCTCGACTGCGGTCCGGAGCCAGAGGCAGAGGCAATAACGCCAACGTAC
GGCTACACCCACCAGAGCCCCTACGCTGCAGTGGCCACGGTGCCAAGTTACTGTTCA**TAA**AGATGATTTG
GTTTGAAGCCATTCTTGAGGATCCTGACAACCAGCAGTGAAACTAGTTTCTACCAGCCTACTAGGCTACC
CTGGACCAAGTGACAAGCTTGAAT

>TaDREB3B **clone 595f1l1-4 1432bp CDS complete, mRNA 79..1157** A-2

GAAGTCGACGCGGCGAAGAAATCAGGCGACAAGATTGCGAACGCTAGATATCTCGACCCGATCCGGGTCG
GGTCGGCC**ATG**ACGGTAGATCGGAAGGACGCTGAGGCGGCGGCGGCGGCCGCACCCTTCGAGATCCC
GGCGCTCCAGCCTGGAA**GAA**CTTGTGGAGCAGAGGAAAGTACCCGGAGTCATGTTCTCGTCAAACCAATA
GGAAAAAGCGACCTCGGAGATCACGTGATGGGCCTAATTCAGTCTCTGAAACGATCAGGCGATGGAAAGA
AGTGAACCAACAACTGGAGCATGATCCACAGGGTGCAAAGAGGGCGAGGAAGCCACCTGCAAAGGGTTCA
AAGAAGGGCTGTATGCAGGGGAAAGGAGGACCTGAGAATACACAATGTGGATTCCGTGGTGTAAGGCAAC
GTACTTGGGGGAAGTGGGTTGCTGAAATTCGGGAGCCAAATCGGGTGAGCAGGCTCTGGTTGGGAACGTT
CCCCACTGCTGAGGATGCTGCCCGTGCTTATGACGAGGCAGCCAGAGCAATGTATGGCGCACTGGCTCGT
ACCAACTTCCCTGTGCATCCTGCACAAGCTCCTGCTGTGGCTGTAGCAGCGGCAATTGAAGGTGTTGTAC
GTGGTGCTTCAGCATCATGTGAGTCTACTACAACATCCACCAACCACTCAGATGTTGCTTCTTCCTTGCC
GCGACAAGCTCAAGCTCTTGAGATTTACTCCCAGCCAGATGTGCTTGAGTCCACCGAATCAGTTGTGCTT
ACTCCTGTTGAGCATTACAGCCATCAAGACAGTGTTCCTGACGCTGGCTCAAGCATTGCAAGGAGCACAT
CCGAAGAGGATGTGTTTGAGCCATTGGAGCCTATTTCCAGTTTGCCGGATGGGGAATCTGACGGTTTTGA
TATAGAAGAATTATTAAGATTGATGGAAGCCGACCCAATTGAAGTTGAGCCGGTCAACGGGGGCTCCTGG
AATGGGGTGGAGATTGGCCAGCAGGAACCTCTCTACCTGGATGGCTTGGACCAAGGCATGCTGGAGGGCA

TGCTGCAGTCTGATTATCCTTACCCAATGTGGATATCAGAGGATCGGGCCATGCACAACCCTGCCTTCCA
TGATGCTGAGATGAGCGAGTTCTTCGAAGGGTTGTGATGAATTACCGGCGGCCAAACCATGTCTATGGTG
TTTGGTCGGCTTGCCCCTTCGGTGTCCGCTGCTGCGCTCCAATGAAGATCAAATGGTGGACCAGATTGGA
TTCCTCTGCAGAACTAATAAGCTCCTAAGCTAGTTTTTTGTGCTTCGTTTGTAGTTCTGTTAGGCATGGG
AACTCTTCTGTTTCGAATGTTTCTTGTTATAAGAAACCTTGATTGTGCATCACGATCTTTGGAAGGTGGA
AAAAGAAAATGTGAAAATGCAAAAAAAAAAAA

>**TaDREB4B** clone 595F1L3-7 1699bp CDS complete, mRNA 139..1234 **A-2**

TGGAATTGTAATACGACTCACTATAGGGCGAATTGAATTTAGCGGCCGCGAATTCGCCCTTGAAGT
CGACGCGGCGAAGAAATCAGGCGACAAGATTGCGAACGCTAGATATCTGGACCCGATCCGGATCGGGCCG
GCCATGACGGTAGATCGGAAGGACGCCGAGGCGGCGGCGGCGGCGACGCCCTTCGAGATCCCCGGCG
CTCCAGCCTGGAAGAACTTGTGGAGCAGAGGAAAGTACCCGGAGTCATGTTCTCGTCAAACCAATAGGAA
AAAGCGACCTCGGAGATCACGTGATGGGCCTAATTCAGTCTCTGAAACGATCAGGCGATGGAAAGAAGTG
AACCAACAACTGGAGCATGATCCACAGGGTGCAAAGAGGGCGAGGAAGCCACCTGCAAAGGGTTCAAAGA
AGGGCTGTATGCAGGGGAAAGGAGGACCTGAGAATACACAATGTGGATTCCGTGGTGTAAGGCAACGTAC
TTGGGGGAAGTGGGTTGCTGAAATTCGGGAGCCAAATCGGGTGAGCAGGCTCTGGTTGGGAACGTTCCCC
ACTGCTGAGGATGCTGCCCGTGCTTATGACGAGGCAGCAAGAGCAATGTATGGCGCCCTGGCTCGTACCA
ACTTCCCTATGCATCCTGCACAAGCTCCTGCTGTGGCTCTACCAGCGGCAATTGAAGGTGTTGTACGTGG
TGCTTCAGCATCATGCGAGTCTACTACAACATCAGCCAACCACTCAGATGTTGCTTCTAACTTGCCGCGA
CAAGCTCAAGCTCTTGAGATTTACTCCCAGCCAGATGTGCTTGAGTCCACCGAATCAGTTGTGCTGGAGT
CTGTCGAGCATTACAGCCATAAAGACACCGTTCCTGATGCTGGCTCAAGCATTGCAAGGAGCACATCCGA
AGAGGATGTGTTCGAGCCATTGGAGCCTATTTCCAGTTTGCCGGATGGGGAAGCAGACGGTTTTGATATA
GAAGAATTACTGAGATTGATGGAAGCCGACCCAATTGAAGTTGAGCCGGTCACTGGGGGCTCCTGGAATT
GTGGAACCAACACTGGCGTGGAGATGGGCCTGCAGGAACCTCTCTACCTGGATGGCTTGGACCAAGGCAT
GCTGGAGGGCATGCTGCAAGCTGATTATCCTTACCCAATGTGGATATCAGAGGATCGGGCCATGCGCAAC
CCTGCCTTCCATGATGCTGAGATGAGCGAGTTCTTCGAAGGGTTGTGATCCCCCTTTGCGGCGGCCAAAC
CATGTCTATGGTGTTTGGTCGGCTTGCCCTTCGGTGTCCGCTGCTGTGCTCCAATGAAGATCAAATGGTG
GACCAGAAGATTGGATTCCTCTGCAGAACTAATAAGCTCCTAGCTAGTTTTTTGTGCTTCGTTTGTAGTT
CTGTTAGGCATGGAAACTCTTCTCTGTTTCGATGTTTCTTGTGATAAGAAACCTTGATTGTGCATCACGA
TCTTTGGAAGGTGGGAAAAGAAAATGTGAAAATGCATTTCCGTGGCAAAAAAAAAAAAAAAAGGCGGCCGC
TCTAGAGTATCCCTCGAGGGGCCCAAGCTTACGCGTACCCAGCTTTCTTGTACAAAGTGGTCCCTATAGT
GAGTCGTATTATAAGCTAGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACTGCTAGCTTGGG
ATCTAAGGGCGAATTCGTTAACC

>**TaDREBP2-2** clone CA662211R1L8-25, 1289bp CDS complete,mRNA 274..1110 **A-2**

C129R1L6-30

CCCCACCCCCGCCCCTGCCCCCGCACCCCACGTCAAAACCAAGGCGGCGGCAGCGGGGCGGGAGAGCGGG
GAGCACCGACCGACACCGGCCGACAGGGCGGGCTGCATGCGGAGCTGAGGCGAGGCGGGGAGAGATCCGG
CGCGGGTGCCACCGCCGGCCGGCCGCGGGAGATCTGGTTGGTGGCGCCGCCCGGATAAGGGAGCGGCCGC
GGAGGCGGCGAGGGGAGAGCAGCCGAAGCGAGAGGAGATCTCTCTCGTCCCTCTTCTCGCTCCATGGAGA
CCGGGGGTAGCAAGCGGGAAGGGGACTGCCCCGGGCAGGAAAGGAAGAAGAAAGTGCGCAGGAGAAGCAC
CGGTCCTGATTCGGTTGCTGAAACCATCAAGAAGTGGAAGGAGGAAAACCAGAAGCTCCAGCAAGAGAAT
GGATCCCGGAAAGCACCGGCCAAGGGTTCCAAGAAAGGGTGCATGGCAGGGAAAGGAGGTCCAGAGAATT
CAAACTGCGCTTACCGCGGTGTGAGGCAGAGGACGTGGGGGAAATGGGTTGCTGAGATCCGTGAGCCCAA
CCGTGGCAATCGGCTGTGGCTTGGTTCATTCCCTACTGCAGTCGAAGCTGCACGTGCATATGATGATGCG
GCAAGGGCAATGTATGGCGCCAAAGCACGTGTCAACTTCTCAGAGCAGTCCCCGGATGCCAACTCTGGTT
GCACGCTGGCACCTCCATTGCTGACGTCTAATGGGGCAACCGCTGCATCACATCCTTCTGATGGGAAGGA
TGAATCGGAGTCTCCTCCTTCTCTTATCTCAAATGGGCCGACAGCTGCGCTGCGTCGGTCTGATGCTAAG
GATGAGTCTGAGTCTGCAGGGACCGTGGCACGTAAGGTGAAGAAAGAAGTGAGCAATGATTTGAGAAGTA
CCCATGAGGAGCACAAGACCCTGGAAGTATCCCAACCAAAAGGGAAGGCTTTACATAAAGAAGCGAACGT
AAGTTATGATTACTTCAACGTCGAGGAAGTTCTTGACATGATAATTGTGGAGTTGAGTGCTGATGTAAAA
ATGGAAGCACATGAAGAGTACCAAGATGGTGATGATGGGTTTAGTCTTTTCTCATATTAGGGTTTTAGCT
ATGAGGGTTGTAGTCATGCGGAGCAATAGGGATAACTTTCATTCTAGCTGCTAGGAAATACTTCAAATTA
TCTGCAACCCGAAGCTCTGTAGTCACTTATGGTTTTCATCTTACTGGAGAGAATAGCTTTATACCATAAG
TCAACGGGTACAAGAAGTTGTCCTGTGCG

>**TaDREBP2-2** clone C129R1L1-42, 1275bp CDS complete, mRNA 315..1079 **A-2**

GTGCCACTCGTGTTGACCCACCCCACCCCACCCCACCCCCCGCCCCTGCCCCCGCACCCCACGTCAAAAC
CAAGGCGGCGGCAGCGGGGTGGGAGAGCGGGGAGCACCGACCGACACCGGCCGACAGCGCGGGCTGCATG
CGGAGCTGAGGCGAGGCGAGGCGAGGCGGGGAGAGATCCGGCGCGGGTGCCACCGCCGGCCGGCCGCGGG
AGATCTGGTTGGTGGCGCCGCCGGGATAAGGGAGAGGCCGCGGAGGCGGCGAGGGGAGAGCAGCCGAAGC
GAGAGGAGATCTCTCTCGTCCCTCTTCTCGCTCG**ATG**GAGACCGGGGGTAGCAAGCGGGAAGGAGACTGC
CCCGGGCAGGAAAGGAAGAAGAAAGTGCGCAGGAGAAGCACTGGTCCTGATTCGGTTGCTGAAACCATCA
AGAAGTGGAAGGAGGAAAACCAGAAGCTCCTGCAAGAGAATGGATCCCGGAAAGCACCGGCCAAGGGTTC
CAAGAAAGGGTGCATGGCAGGGAAAGGAGGTCCAGAGAATTCAAACTGCGCTTACCGCGGTGTGAGGCAG
AGGACGTGGGGCAAATGGGTTGCTGAGATCCGTGAGCCCAACCGTGGCAACCGGCTGTGGCTTGGTTCAT
TCCCTACCGCAGTCGAAGCTGCACGTGCATATGATGATGCGGCAAGGGCAATGTATGGCGCCAAAGCACG
TGTCAACTTCTCAGAGCAGTCCCCAGATGCCAGCTCTGGTTGCACGCTGGCACCTCCATTGCTGATGTCT
AATGGGGCAACTGCCGCATCACATCCTTCTGATGGGAAGGATGAGTCTGAGTCTGCAGGGACCGTGGCAC
ATAAGGTGAAAAAAGAAGTGAGCAATGATTTGAGAAGTACCCATGAGGAGCACAAGACCCTGGAAGTATC
CCAACCAAAAGGGAAGGCTTTACATAAAGAAGCGAATGTAAGTTATGATTACTTCAACGTCGAGGAAGTT
CTTGACATGATAATTGTGGAATTGAGTGCTGATGTAAAAATGGAAGCACATGAAGAGTACCAAGATGGTG
ATGATGGGTTTAGTCTTTTCTCATAT**TAG**GGTTTTAGCTATGAGGGTTGTAGTCATGCGGAGCAATAGGG
ATAACTTTCATTCTAGCTGCTAGGAAATACTTCAAATTATCTGCAACCCGAAGCTTTGTAGTCACTTATG
GTTTTCATCTTACTGGAGAGAATAGCTTTATACCATAAGTCAACGGGTACAAGAAGTTGTCCGTGCG

>**TaDREBP4A** clone C129R1L1-44, 972bp CDS complete, mRNA 48..746 **A-3**

GAATTCCCGGGATATCGTCGACCCACGCGTCCGATCTCATCTCATCA**ATG**GAGCAGGAGGTGGTTGCGGG
GATGAAGCAGAAGAAGTGCTGCCCGCTCCGGCGGTCGCGCAAGGGCTGCATGAAGGGCAAGGGCGGCCCG
GACAACCAGCAGTGCCCCTTCCGCGGCGTCCGCCAGCGCACCTGGGGCAAGTGGGTCGCCGAGATCCGCG
AGCCCAACCGCGGCGCCCGCCTCTGGCTCGGCACCTTCGCCACCGCGCTAGACGCCGCGCGCGCCTACGA
CGCCGCGGCCAGGGCGCTCTACGGCGACTGCGCCCGCCTCAACCTCTCGGCGTCGCCGTCCCAGATGCAG
CAGCACCCTCCAGCTCAAGGCAGTGGTGGCGCCAATGGTAATTCGGCGCCAGGGACACCGTGCTGCTCCT
CCAACAACTCCAACTCCAGCGCGTCGACCCCGACCGGGACTCCCACGGACATGGACTGCAGCGCCTGGAT
GCAGCCGTCTTACTGTTACAGCACGGCGGAGGCGCCGGAGGACTTCGAGGCGTACGTGACGCGACTGCCC
AAGGCGGAGGACTTCGGACTGGAGGGGTTCCAGGAGGTGCCCCTGGAAGTGCTGGCCGAAGCCGGAGGAG
GGGTCAGCATATGGGACCTCTCCATCGCCCCTGACATGGCTGCAGCAGCCGCTTCTTCAGCTGCTGCCAG
TGCCTGCACTGTCCCCCAACAGCCGCTGCAGCAACCCAGCTGC**TGA**GGTCCAGATCGACGTCGACATAGA
AAGTAAGGAGTGGCCGTGATATACAGTGCATATACAGCATGCGGCATGCTGGTCACTGGCCGTGGTGCAT
GCAGCTAGATAGCTACGTAGTTAGTACTAGTATTTGCTTTTTCTCCGTCGTCGCAGTAGCATACTAGCAT
TAATATTGGTGGGCAGGTGAGAGTGCAAGTGTGTGTATGATATAATCGTGCGTGGGAGCAGG

>**TaCBF7-1** clone WEF77D03, 964bp CDS complete,mRNA 89-**916 A-4**

GGAGGTTGTCCCGCTCCGGATTCCCGTGATATTGTCGCCCCACGCGTCCGCACCTAGCTACAGGCCTTCA
GATCCCTCCTAACGAGCC**ATG**GAGGCAGACGCAAGCCATGCACCCACCACCTCTTCCTCCTCCGTCTCGT
CGTCGACATTGTCCACCTCCTCCTCCTCCTCCCTCGCCACCAGTGTGCTAGAGCTCCCCAAGAACCC
CAAGCCCAAGCACCCGAAGAAACGCAAGAGAGCCGCCGCCGACCAAGAAACGGACGCCGCCGCGGCCACC
ATTGGCGCACGAGGAGACGAGAGCAGCTGCTGCAGCACCGACGAAGACAACGCCGCGAGCGTCAAGGCAG
CCGTGTCCAAGAGCGGCTTCAAGCACCCGTCGTACCGCGGCGTGCGGCGCCGGAGCTGGGGTAAGTGGGT
GTCCGAGATCCGTGAGCCGCGCAAGAAGTCGCGCATCTGGCTCGGCACCTTCCCCACCGCGGAGATGGCG
GCGCGAGCCCACGACGTGGCCGCGCTCGCCATCAAGGGCCGCTCCGCGCACTTCAACTTCCCGGAGCTCG
CCCACGAGCTGCCCCGCCCGGACTCCACGTCGCCCGCGGACATCCAGGCCGCCGCCGCGAAGGCCGCCGC
CACCGCCGCCGTGCAGTGCGAGGCCGAGCCGGAGCACGAGCACGAGCCCGAGACGCCGTCGTCGTCCGGC
GCCGTTTCGGAGACACCGGAGGCTGCAGCCTGCACCGAAGCGGTGCCCGCGGACAGGGGCGAGGTCGACA
ATGCGCTTTTCGACCTGCCCGACCTTCTTCTGGACCTGAGGGATGGGCTCTTCTGGTCGCCGGTCTGGCC
GGTGGCGCTGGCCGCCGAGGAGTACGACGGGGGCTGCTGCGTTGGGCTCAGTGAGCCTCTCTTGTGGGCC
GAG**TAG**GAATATATTCTTCAGTTTTTACCGTTCAAAAAAAAAAAAAAAAAAAA

>**TaCBF7-2** 902bp CDS complete,mRNA 54-**887 A-4**

GAATTCGCCCTTGTCCACCTAGCTACAGGCCCTCAGATCCCTGCTACCGAGCC**ATG**GAGGCAGACGCGAG
CCATGCACCCACCACCTCTTCCTCCTCCGTCTCCTCGTCGACATTGTCCACCTCCTCCTCCTCCTCCTCC
CTCGCCACCAGCGCGCTAGAGCTCCCCAAGAGCCCCAAGCCCAAGCACCCCAAGAAGCGCAAGAGAGCCG
GCTCCGACCGAGAGCAGGACGCCGCCGCGAACACCAATGGTGTCCGTGGAGACGAGAGCAGCTGCTGCAG
CACCGACGAAGACTTCGCCGCCGCGAGCGTCAAGGCGGCCGTGCCCAAGAGCGGCTTCAAGCACCCGTCG
TACCGCGGGGTGCGGCGCCGGAGCTGGGGCAAGTGGGTGTCCGAGATCCGCGAGCCGCGCAAGAAGTCGC
GCATCTGGCTCGGCACCTTCCCGACCGCCGAGATGGCGGCGCGCGCCCACGACGTGGCCGCGCTCGCCAT
CAAGGGCCGCTCCGCGCACCTCAACTTCCCGGAGCTCGCGCACGAGCTGCCCCGCCCGGACTCCACGTCG
CCCGCGGACATCCAGGCCGCCGCCGCGAAGGCCGCCGCCACCGCCGCCGTGCAGTGCGAGGCCGAGCCCG
AGCACGAGCACGAGCCGGAGACGCCGTCGTCCTCCTCCGGCGCCGTCTCGGAGACACCGGAGGCCGCAGC
CTGCGCCGAAGCGGCGCCCGCGGACAGCGGCGAGGTCGACAATGCGCTGTTCGACCTGCCCGACCTCCTT
CTGGACCTGAGGGACGGGCTCTTCTGGTCGCCGGTCTGGCCGGCGGCGCTGGCCGCCGAGGAGTACGACG
GCGGCTGCTGCGTCGGGCTCAGTGAGCCTCTCCTGTGGGCCGAG**TAG**GAAAAGGGCGAATTC

>**TaDREBP-1** clone C79R1L8-47, 1275 bp CDS complete,mRNA 141..872 **A-5**

CCAGCTCCGCCCCGCTGCACACCAAAAGCTACCCCTCCTTCGGGCACGCGGGCGGCCGGCAGGGGAGAGA
GAGGGCAGCCGAGGCGGCTAGCCAGCGCAGGAGCTTAGGTAGGTTGTAGGCAGGCAGCCTCGCTTCGCTC
ATGCAGCAGGGCGAGTACCGCTCGTCGTCTTCCAGCGAGGGCTCCGCGGGGTCGGCTGCTGCGGCTGCGG
CGGCCGCGGCCATGGCGCCCCTGGCGGCTGCGGCCGCGGCGGTGGCGGCCAAGGAGGAGCACAACGTGAC
GGTGGCCGTGGCGCCGCCCATGCCCATGGCGATGGCGATGCCGCTGCAGCAGCAGCAGCCGCGGAAGCAG
TACCGCGGCGTGCGCATGCGCAAGTGGGGCAAGTGGGTGGCGGAGATCCGCGAGCCGCACAAGCGCACGC
GCATCTGGCTGGGCTCCTACGCCACCCCCGTCGCCGCCGCGCGCGCCTACGACACGGCCGTCTTCTACCT
GCGCGGCCGGTCGGCGCGCCTCAACTTCCCCGACGAGATCTCCGCGCTGGCGCTGTCGTCGCCCGAGGCC
GCCGAGGCCGGCGGAAGGGAGGAGCCGGGCGGCGGCGGCGGCGCGCTGTCGGCCGCGTCGATCCGGAAGA
AGGCCATCGAGGTCGGGTCCCGCGTGGACGCGCTGCAGACCGGCATGACCACCATGGTCGCCGCGCCCGC
GCACCACCGCGAGCGGCAGCGGCTCCACCACCACCACCACCACGCGGAGCCGCACGGCGAGGAGCTGCAC
CGCCACGTGAAGCAGCAGCGGACGGCGTGGAACGGGCGCGCCAAGAACCCGGATCTCAACCAGGCGCCGA
GCCCGGACACCTCCGACGCCGAGGCCGAG**TGA**AGCAGCTAGGCATCAAAGAGAAGCAGCGGCTTCCAGTC
CAGCCATCCACAACCCGCCGCCCATGCAGCTAGCTAGTCTCCGGAGAGCCGCCAGCCGGCGTGTGATCAA
TCAGGTTCAGCACAAGCAGCTCAATCCAGTCCACCCACCCACCTACTGTTCCGCCTGTCTCCCTAGCAAG
CTCCGCCGGGTGCGCGTCGGAGAGGGCGGCGAGGTGCCGGTGGTCGGTGGGTCGGTGGCCGGATTTTTCA
ATACGTCGTTCTGGATAAGAACAAGGGAGGGGGAGGGGAGGCGGGGGACGCCACAAGTGGCGGTCTTTCC
AAATGTCAAAAAAGACAGCTGTAACAGTGATAAAACAATCGCCACCATCTTCTTTTCTCTCTCTCTCTCG
CCGTCTCTTGCTGGC

>**TaDREBP-2** clone c82r1L1-5 , 958 bp CDS complete,mRNA 71.. 739 **A-5**

GAATTCCCGGGATATCGTCGACCCACGCGTCCGATCTATCTCATCAATGGAGCAGGAGGTGGTTGCGGGG
ATGAAGCAGAAGAAGTGCTGCCCGCTCCGGCGGTCGCGCAAGGGCTGCATGAAGGGCAAGGGCGGCCCGG
ACAACCAGCAGTGCCCCTTTCGCGGCGTCCGCCAGCGCACCTGGGGCAAGTGGGTCGCCGAGATCCGCGA
GCCCAACCGCGGCGCCAGCCTCTGGCTCGGCACCTTCGCCACCGCGCTCGACGCCGCCCGCGCATACGAC
GCCGCGGCCAGGGCGCTCTACGGCGACTGCGCCCGCCTCAACCTCTCGGCGTCGCCGTCCCAGATGCAGC
ACCCTCCAGCGCAAGGCAGTGGCGCGAACGGTAATTCGGCGCCGGGGACGCCGTGCTGCTCCTCCAACAA
CTCCAACTCCAGCGCGTCGAACCCGACCGGGACTCCCACGGAAATGGACGGCAGGGCCTGGATGCAGCCG
TCTTACTGGTACAGCACGGCGGAGGCCCCCGAGAACTTCCAGGCGTACGTGACGCGACTGCCCAAGGCGG
AGGACTTCGGACTGGAGGGGTTCCAGGAGGTGCCTCTGGAGGTGCTGGCCGAAGCCGGGGGAGGGGTCAG
CATCTGGGACCTCAGCATCGCCCCTGACATGGCAGTAGCAGCCGCTTCTTCAGCTGCTGCCAGTGCCTAC
ACTGTCCCCCAACAGCCGCTGCAGCAACCCAGCTGC**TGA**GGTGCAGATCGACGTCGACATAGAAAGTAAG
GAGTTGCCGTGATATATAGTGCATATACAGCATGCGGCATGCTGGTCACTGGCCGTGTTGCATGCAGCTA
GCTACGTAGTTGGTAGTATCTGCTTTTTCTCCATCGTCGCAGTAGCATACTAGCATTAATATTGGTGGGC
AGGTGAGCGTGCAAGTGTGTGTATGATATAATCGTGCGTGGGAGCAGG

>**TaDREBP-3** clone c82r1L1-6 , 967 bp CDS complete,mRNA 43.. 740 **A-5**

GAATTCCCGGGATATCGTCGACCCACGCGTCCGATCTCATCA**ATG**GAGCAGGAGGTGGTTGCGGGGATGA
AGCAGAAGAAGTGCTGCCCGCTCCGGCGGTCGCGCAAGGGCTGCATGAAGGGCAAGGGCGGCCCGGACAA
CCAGCAGTGCCCCTTTCGCGGCGTCCGCCAGCGCACCTGGGGCAAGTGGGTCGCCGAGATCCGCGAGCCC

AACCGCGGCGCCCGCCTCTGGCTCGGCACCTTCGCAACCGCGCTCGACGCCGCCCGCGCATACGACGCCG
CGGCCAGGGCGCTCTACGGCGACTGCGCCCCGCCTCAACCTCTCGGCGTCGCCGTCCCCAGATGCAGCAC
CCTCCCAGCGCAAGGCAGTGGCGCGAACGGTAATTCGGCGCCCGGGGACGCCGTGCTGCTCCTCCCAACA
ACTCCAACTCCAGCGCGTCGACCCCGACCGGGGACTCCCACGGACATGGACTGCAGTGCCTGGATGCAGC
CGTCTTACTGTTACAGCACGGCGGAGGCGCCGGAAGACTTCGAGGCGTACGTGACGCGACTGCCCAAGGC
GGAGGACTTCGGACTGGAGGGGTTCCAGGAGGTGCCCCTGGAAGTGCTGGCCGAAGCCGGAGGAAGGGTC
AGCATATGGGACCTCTCCATCGCCCCTGACATGGCTGCAGCAGCCGCTTCTTCAGCTGCTGCCAGTGCCT
GCACTGTCCCCCAACAGCCGCTGCAGCAACCCAGCTGC**TGA**GGTCCAGATCGACGTCGACATAGAAAGTA
AGGAGTGGCCGTGATATACAGTGCATATACAGCATGCGGCATGCTGGTCACTGGCCGTGGTGCATGCAGC
TAGATAGCTACGTAGTTAGTACTAGTATTTGCTTTTTCTCCGTCGTCGCAGTAGCATACTAGCATTAATA
TTGGTGGGCAGGTGAGAGTGCAAGTGTGTGTATGATATAATCGTGCGTGGGAGCAGG

>**TaDREBP-4** clone C42F1L1-A(39), 1342bp CDS complete,mRNA 266..985 or **A-5**

CCCGGTTTATGATTATCCCAACCAAACTCTAACCATACCCCCACCCACAAAATCCCAATACAAATACCAC
CGCCCAGCAAATAGATAGATAGCCCGGACCGGAGGGAGGGAAGAGAGAGTCCAGCTCCGCTGCCGCAC
ACCAAAAGCTACTCCTCCTTTGGGCACGCGGCCGGCAGGGAGGGGAGAGGGGAGTCGAGGCGGCGGCCGG
CGGCTAGCCAGCATAGGAGCATAGGTAGTTGGTAGGCAGCTCAGCTTCGCTC**ATG**CAGCAGGGCGAGTAC
CGCTCGTCGTCTTCCAGCGAGGGCTCCGCGGGGTCGGCGTCGGCTGCGGCGGCGGCGATGGCGCCTCTGG
CGGCTGCGGTCGCGGCGGTGGCGGCCAAGGAGGAGCACAACGTGACGGTGGCCGTGGCGCCGCCCATGCC
CATGGCGATGGCGATGCCGCTGCAGCAGCAGCAGCCGCGGAAGCAGTACCGCGGCGTGCGCATGCGCAAG
TGGGGCAAGTGGGTGGCGGAGATCCGCGAGCCGCACAAGCGGACGCGCATCTGGCTGGGGTCCTACGCCA
CCCCCGTCGCCGCCGCGCGCGCCTACGACACGGCCGTCTTCTACCTGCGCGGCAGGTCGGCGCGCCTCAA
CTTCCCCGACGAGATCTCCGCGCTGGCGCTGTCGTCGCCCGAGGCCGCCGCCGAGGCCGGAGGGGAGGAG
CCGGGCGACGGCGGCGGCGCGCTGTCGGCCGCGTCGATCCGGAAGAAGGCCATCGAGGTCGGGTCCCGCG
TGGACGCGCTCCAGACCGGCATGACCACCATGGTCGCCGCGCCCGCGCACCACCGGGAGCGGCAGCGGCT
CCACCAACACCACCACCACGCGGAGCCGCACGCTGAGGAGCTGCACCGCCACGTGAAGCAGCAGCGGACG
GCGTGGAACGGGCGCGCCAAGAACCCGGATCTCAACCAGGCGCCGAGCCCGGACACCTCCGACGCCGAG**T**
GAAGCAGCCAAGCAGCGAGAAGAAGCAGCGGCTGCTTCCAGTCATCCACAACCCGTCGACCATGCAGCTA
GTCTCCGGAGAGCGGCCAGCCGGCGTATGATCAATCAGATTCAGCACAACCAACTCAATCGTGTCCACCT
GTCCCCCTAGCGAGCTCCGCCGCGTCGGTGAGGGCGGCGAGGTGCCGGTGGTCGGTGGGTCGGTGGCCGG
ATTTTTCAATACGTCGTTCTGGATAAGAACAAGGGAGGGAGGGGAAGCGGGGGACGCCACAAGTGGCGGT
CTTTCCAAATGTCAAAAAAAACAGCTGTAACAGTTGATAAAACAATCGCCACCATCTTCTTTTTAAAAAA
AAAAAAAAA

>**TaDREBP-5** clone C116R1L1-28, 874bp CDS complete,mRNA 179..718 **A-5**

CGACTGGAGCACGAGGACACTGACATGGACTGAAGGAGTAGAAAGTACCAGTGGAGAGAGAGGCCGGGCC
GACACACACACAGCGGCAGAGCGCCCACGCCTCAGGCAGGCGCCGCACCAGCAACCAACCAGCCCATCCC
CTAGGAGCCGCCGCCGAATTGGGTGCGCGGGCGCCGCC**ATG**GAGAGGGAAGCCACGGTCGCCTTCTACCC
GCCGGCGCCGGCGCCGCAGCAGCAGCGCCCCACGGCCGCGCAGCCGCTCTCGGCCCGCGTGCCTGCCGGC
GGCGTCCACGCGGGCCGGGGAGGAGGGGGCAGGCAGTACCGCGGGGTGCGGATGCGCAAGTGGGGCAAGT
GGGTGGCCGAGATCCGGGAGCCCAACAAGCGGTCGCGGATCTGGCTCGGCTCCTACGCCACCGCCGTCGC
CGCCGCGCGCGCCTACGACACCGCCGTCTTCTACCTCCGCGGCCGCTCCGCGCGCCTCAACTTCCCCGAC
CAGCTCCTCGACGGCACTGCCCCCGCCGCCGCGCCCGGGGACCTCACCGCCGCCGCCATCCGCAAGAAGG
CCGCCGAGGTCGGCGCGCGCGTCGACGCGCTCCACTCCGGCGGCATCGGGGTCGGCATGGGCGCCCCTGC
GGCACCGCCGTCGCCGTCCCAGCGCCGCAGGGCCAAGAACCCCGACCTCAACCGGGAGCCCACCCCCGAC
ACCGACGACGACGAG**TGA**ACTGCTGTAAACCCCCGCCAATTCTTCAGTTCTTCTCCGACCACCATCACCA
CCTACATACCTAGTACTAGTGCGTCTTAATTAGTGGTTAGCAACCAGCAGCGAGGTAGCTAGCGGTAGCA
GAGCATCAACCATCTCAGTGAGCAAGCAGATACT

>**TaDREBP3-1** clone c20r111-a, 1882bp CDS complete,mRNA 643..1698 **A-5**

GAATTCCCGGGATATCGTCGACCCACGCGTCCGCGACTGGAGCACGAGGACACTGACATGGACTGAAGGA
GTAGAAAAGAGAGACGCCTCAACCCCCACCATCCACAAGAACAAGCAGCCACAATCCCATTCCACTCCAC
GGTGAGGTTGAGCGATCTGACGAGCCCTTCAGCAGCAGCGACGACGGCGGCAGGTTTGCTCTCTGAACTT
CTCTGATCTTTTTAGCGGCAGCCAGCAGGAATACACAAGAAACCGATCGAGGTCCCGTTCAAGGGGTTCT
TCTTCTGGGATTATTAGATCAGAAGGGGCGCCCTCCTCTTGCTTGGGATTCTCTTGATTTCTTGCCAACC

TCACCACTCTCCTAGGAGGAACAGAGAGATCCAGAGAACCTTTTTTCTCTCCTGGATTCTTGCCCTATTC
AGCCTTGTTTTGCTTCGGTACGTCCGGAAAGGGTTCTTTTGCAATACCCTTGTAGTTCGTCCGATTTATT
CTTTTCTTCCCCTCGTTTCTCTAATTTTATTGAGGAACAGGGCCTTTTTTCTCTCTCTTTCTGCTTGCAC
GAGAAAGCTCCTTTTTTTTCATTCCGTTTTTGGTTCGGCCTTCTCTACATCACCTTCTTCTTCTACCTCT
GCCCAAGCCCTC**ATG**GCCGCTGCCATAGACATGTACAAGTACAACACCAGCACGCACCAGATCGGCTCTG
CCGCCTCTGCTTCGGATCAGGAGCTCATGAAAGCACTCGAACCTTTTATCACGATTGCTTCCTCCTCTCC
CTACCCCTACCAGTACTACTCTTCTCCCTCCATGACCCAAGATTCATACACGGCCACCCCATCATCATCG
TACGCCTCGTTCGCGACCTCTCCTCTACCCACCACCGCACCCACCTCGCCGTCTTTCTCGCAGCTTCCGC
CGCTCTACTCCTCGCAGTATTCCACCGCTTCGGGCATGAACGGATCCATGAGGCTGGCCCAGCTCGGCCC
GGCCCAGATCCAGCAGATCCAGGCCCAGTTCTTCGTCCAGCAGCAGCAGCAGCAGAGGGGCCTGGCTGGC
TCCTTCCTTGGGCCGCGCGCGCAGCCGATGAAGCAGTCCGGGTCGCCGCCGCGCGGCGTCCGCCGCGGCGC
TGGCCCTGGCCGGAGTGGCGCCTGCGCAGTCCAAGCTGTACCGCGGCGTGCGGCAGCGTCACTGGGGCAA
GTGGGTGGCGGAGATCCGCCTCCCCAAGAACCGCACGAGGCTGTGGCTCGGCACCTTCGACACCGCCGAG
GACGCCGCGCTCGCCTACGACAAGGCCGCCTTCCGCCTCCGCGGGGACCTCGCCCGCCTCAACTTCCCGT
CGCTCCGACGCGGCGGCGCCCACCTGGCCGGCCCGCTCCACGCCTCCGTCGACGCCAAGCTCACCGCCAT
CTGCGAGTCCCTCGCCGCGCCCTCGTCCAAGAACTCGGCCGAGGCGGAGCCGGAGTCCCCCAAGTGCTCG
GCGTCGACGGAGGGTGAGGACTCGGCGTCCGCCGGGTCCCCCCCCCGCCGCCCACGCCGCCGGTCCCGGA
GATGGAGAAGCTGGACTTCACGGAGGCGCCGTGGGACGAGTCGGAGACCTTCCACCTGCGCAAGTACCCG
TCCGTGGAGATCGACTGGGACTCCATCCTGTCGTGAACAGCAAGCAGCTGCTACTACTACTAGTACAGTC
TTCGTAGCTATGATG**TAA**TTTCTCTTTCGCTCGGATCGGCGGCTGCTCTGGCCCGACGGCATTTTAGACG
TCGGCCATGGCTGCTGCGAGTAGCAGTAATTAACTAGTAGCAGTAAGTAGCCAGTGGTGTTTAGTAAGAT
CGTTGCTACTACTACGTGGTGTAATTGATCTCCTGGTCGACCTGCCGGCAAAGGGCGAATTC

>**TaDREBP3-2** clone C106R1L7-10, 1720bp CDS complete,mRNA 537..1532 **A-5**

CCACGTAGTAGCAACACGCAGCCACATCCACATCCCATTCCACTCCACGGTGAGGTTGAGCGATCCGACG
AGCCCTTCAGCAGCAACGACGGCAGGTTTGTTCTCTGAACTTCTCTGGTCTTTTTAGCGGCAGCCAGCAG
GAACACACAAGAAATTCAGGTCCCGTTCAAGGGGTCCTTCTTCTGGGATTATTAGATCGGAAGAAGCGCC
CTCCTCTTGCTCGGGATTCTCCTGATTTCTTGCTAACCTCACTACTCTCCTAGGAGGAACAGAGAGATCC
AGAGAACCTTTTTTCTCTCCTGGATTCTTGCCTTTTTTCAGCCTTGTTTTGCTCCGGTACGGCGAGAAAG
GGTTCTTTTGCAATACCCTTCTAGTTCGTCCGATTTGTTTCCTCTTCTTCCCCTCGTTTCTCTAATTTTA
TTGAGGAACAGGGCCTTTTTTTTCTTCTGATTTTGCACCAGAAAGCTCCTTTTTTTTTCATTTCGGTTTTGG
TTCGGCCTTCTCTACATCACCTTCTTCTACCTCTGCCCAAGCCCTC**ATG**GCCGCTGCCATAGACATGTAC
AAGTACAACACCAGCACGCACCAGATCGGCTCCGCCGCCTCTGCCTCGGATCAGGAGCTCATGAAAGCAC
TCGAACCTTTTATCACGATTGCTTCCTCCTCTTCCTCTCACTACCCCTACCACTACTACTCCTCCCCCTC
CATGACCCAAGATTCGTACATGGCCACCCCATCGTCGTCGTCCTACGCCTCCTCGTTCGCAGCCTCTCCT
CTGCCCACCACCGCCCCCGCCTCGCCGTCTTTCTCCCAGCTTCCGCCGCTCTACTCCTCGCAGTATGCCG
TGAACGGATCCATGGGGCTGGCCCAGCTCGGCCCGGCCCAGATCCAGCAGATCCAGGCCCAGTTCTTCGT
CCAGCAGCAGCAGCAGCAGCAGAGGGGCCTAGCTGGTGGCTCCTTCCTCGGCCCGCGCGCGCTGCCGATG
AAGCAGTCCGGGTCGCCGCCGCGCGCGTCCGCCGCCGCGCTGGCGCTGGCCGGAGTGGCGCCCGCGCAGT
CCAAGCTGTACCGCGGCGTGCGGCAGCGGCACTGGGGCAAGTGGGTGGCGGAGATCCGCCTCCCCAAGAA
CCGCACCAGGCTGTGGCTCGGCACCTTCGACACCGCCGAGGACGCCGCGCTCGCCTACGACAAGGCCGCC
TTCCGCCTCCGCGGCGACCTCGCCCGCCTCAACTTCCCGTCGCTCCGGCGCGGCGGCGCCCACCTGGCCG
GCCCGCTCCACGCCTCCGTCGACGCCAAGCTCACCGCCATCTGCGAGTCCCTCGCCGCGCCCTCGTCCAA
GAACTCGGAGCCGGAGTCCCCCAAGTGCTCGGCGTCGACGGAGGGCGAGGACTCGGCGTCCGCCGGGTCC
ACGCCGCCGGTCCCGGAGATGGAGAAGCTGGACTTCACGGAGGCGCCGTGGGACGAGTCGGAGACCTTCC
ACCTGCGCAAGTACCCGTCCGTGGAGATCGACTGGGACTCCATCCTGTCG**TGA**GCAGCAAGCAGCTGCTG
CTACTACTAGTACAGTCTTCGTTAAGCTCCGTAGCTATGATGTAATCTCTCCTTCGATCGAATCGGCGGC
TGCTCTGGCCCGACGGCATTTTAGACGTCGGCCATGGCTGCTGCGAGTAGCAGTAATTAACTACTACTAG
TAGTAATTAGCCAGTGGTGTTTAGTATGCTACTACGTGGT

>**TaDREB2-1** clone WEF071-J23, 1845bp CDS complete,mRNA 536..1549 **A-6**

CACGAGCAACACGCAGCCACAATCCCATCCCACTCCACGGTGAGGTTGAGCGATCTGACGAGCCCTTGAG
CAGCAGCAGCGACGACGGCGGCAGGTTTGCTCTCTGAACTTCTCTGGTCTTTTTAGCGGCAGCCAGCAGG
AACACACAGGAAACCAAGGTCCCGTTCAAGGGGTTCGCTTCTTCTGGGATTATTAGATCGGAAGAGGCGC
TCTCCTCTTGCTCGGGATCTTCTTGATTTCTTGCCAACCTCACCACTCTCCTAGGAGGAACAGAGAGATC
CAGAGAACCTTTTTTCTCTCCTGGATTCTTGCCCTTTTTCAGCCTTGTTTTGCTTCGGTACGGCCGGAAA

GGGTTCTTTTGCAATACCCTTCTAGTTCGTCCGATTTATTCCCCTCGTTTCTCTAATTTTATTGAGGAAC
AGGGCCCTTTTTTTTCTTTCTTTCTGCTTGCACGAGAAAGCTTCCTTTTTTCCATTCCGTTTTTGGTTCG
GCCTTCTCGACATCACCTTCTTCTTCTACCTCTGTCCAAGCCCTC**ATG**GCCGCTGCCATAGACATGTACA
AGTACAACACCAGCACGCACCAGATCGGCTCCGCCGCCTCTGCTTCGGATCAGGAGCTCATGAAAGCACT
CGAACCTTTTATCACGATTGCTTCCTCCTCTTCCTCTCACTACCCGTACCAGTACTACTCTTCTCCCTCC
ATGACCCAAAATTCATACATGGCCACCCCATCGTCGTCCTACGCCTCCTCGTTCGCAGTCTCTCCTCTGC
CCACCACCGCGCCCGCCTCGCCGTCTTTCTCGCAGCTTCCGCCGCTTTACTCCTCGCAGTATGCCGCTTC
GGGCATGAACGGATCCATGGGGCTGGCCCAGCTCGGCCCGGCCCAGATCCAGCAGATCCAGGCCCAGTTC
TTCGTCCAGCAGCAGCAGCAGCAGAGGGGCCTGGCTGGTGGCTCGTTCCTTGGGCCGCGCGCGCAGCCGA
TGAAGCAGTCCGGGTCGCCGCCGCGCGCGTCCGCCGCGGCGCTGGCGCTGGCCGGAGTGGCGCCCGCGCA
GTCCAAGCTGTACCGCGGCGTGCGGCAGCGCCACTGGGGCAAGTGGGTGGCGGAGATCCGCCTCCCCAAG
AACCGCACGAGGCTGTGGTTTGGCCCCTTTGGACCCCGCCGGGGGACGCCGGCGTTCGCTTACGACAAGG
CCGCCTTTCCGCTTTCGGGGGGGGACTTCGCCCGGCTTCAACTTCCCGTCGGTTCCGCCGGGGGGGGCGC
CCAACTTGGCCGGGCCCGTTTCCAGGCTTCCGTGGACGCCAAGTTCACCGCCATTTGCGAGTCCCTCGCC
GCGCCTTTGTCCAAGAACTCGGAGCCGGAGTCCCCCAAGTGCTCGGCGTCGACGGAGGGCGAGGACTCGG
CGTCCGCGGGGTCCCCGCCGCCGCCCACGCCGCCCGTCCGGGAGATGGAGAAGCTGGACTTCACGGAGGC
GCCGTGGGACGAGTCGGAGACCTTCCACCTGCGCAAGTACCCGTCCGTGGAGATCGACTGGGACTCCATC
CTGTCG**TGA**ACAGCAAGCAGCTGCTACTACTAGTACAGTCTTCGTTAAGCTCCGTAGCTATGATGTAATC
TCTCCTTGGATCGAATCGGCGGCTGCTCTGGCCCGACGGCATTTTAGACGTCGGCCATGGCTGCTGCGCG
TAGCAGTAATTAACTAGTAGCAGTAAGTAAGTAGCCAGTGGTGTTTAGTAAGATCGTTGCTACTACTACG
TGGTGTAATTGATCTCCTGGTCGACCTGCCGGCAGTTTTTTTCACGGCAAGGCGGCCAGTCGAGAGGTGT
AATCATGTTTACCCGTTGAAAGAAT

>**TaDREB2-2** clone C63R1L1-11 1679bp CDS complete, mRNA 498. 1349 **A-6**

CCCCGCCTCCATTTCGCCCACTCCGACCTTCGTCCACACACCTTCCACCAGCCACCGAAAAAGCCTCAAT
CCCAGCAACCACAGACCACTCACGTGAGTTCGAGCGAGGTAAGCATAGATCCGAGCGAGTTCTTGCGTCG
CCAAGGTAAGGTTCTCCGAGTCTCAAGCCCTCTTCCTCCCGCCCCCCTATTCGTCTCCTTGTTCACCCGG
CAGGCAGCGAGGTCTCCAGGACGTTTTAGATCCCGGGGCTCCAGATCTGTTCTTTTTTCTTCTCTTTTAG
TTGTAGTAATCGGCCAGGCGAGGTTCTTTTTCCACACCTCCTTGCTCGTTGTGTCTAGATCTACCACCAA
CGCCATCTCTTTTTGCCGTGGCTGTTGACCTTGTTGATTTTCTCTGCCTGTTCTAGGTTTTGGTTTGGAA
ACCAGGCAGATTTCTTTCTTGTTCATCTTTCTAAGAAAACAAAAATCCCCAAAAAAAAGTTCGTCCTAGT
TCCTTAG**ATG**GCTGCAGCTATAGATCTGTCCGGGGAGGATCTGGTGAGAGCACTCGAGCCTTTCATCCGA
GAAGCCTCTGCCCCCCTTCCACTCCACTCCCATCTTAGTCCCACCTCGCCATTCTCCTTCCCCCACGCCG
CCTACAGTGGGTACCCGTACGGGGTGCAGGCACAGGCCCAGACCGAGCTCAGCCCGGCCCAGATGCACTA
CATCCAGGCACGCCTCCACCTCCAGCGCCAGACTGGCCAGCCGGGCCAGCTCGGACCGCGGTCCCAGCCC
ATGAAGCCCGCTTCGGTGGCAGCGGCGACCCCGCCGCGGCCGCAGAAGCTCTACCGTGGCGTGCGGCAGC
GCCACTGGGGCAAGTGGGTGGCGGAGATCCGTCTCCCCCGCAACCGCACCCGCCTCTGGCTCGGCACCTT
CGACACCGCCGAGGAGGCGGCCCTCGCCTACGACCAGGCCGCCTACCGCCTCCGCGGCGATGCAGCTCGC
CTCAACTTCCCCGACAACGCCGCCTCTCGCGGCCCGCTCCATGCCTCCGTTGACGCCAAGCTCCAGACCC
TCTGCCAGAACATCACCGCCTCCAAGAACGGCAAGAAGTCCGCCTCCGTCTCCGCGTCCACCGCCGCAGC
CACGTCGTCCACCCCCACCAGCAACTGCTCCTCGCCGTCCTCCGACGAGGCGTCGTCCTCGCTGGAGTCC
GCGGAGTCGTCACCGTCACCCGCCACCACCACCGCAGCAGAGGTTCCAGAGATGCAGCAGCTCGACTTCA
GCGAGGCGCCATGGGACGAGGCAGCCTGCTTCGCCCTCACCAAGTACCCGTCGTACGAGATCGACTGGGA
CTCGCTTCTCGCCACC**TAA**TAGCACCCAGTTTGTGTTCGTCAGCTACTACTACTACCGTCTTTAGCGTTT
CATGATGCTAGGTTAATGGGTCGCCGCGATGCAGATGGCATTTTAGACATTCTGCACGGGCTTTTAGCGG
ATTAGCTCTAAGTCTCTAATCCTTGTTCATTGTGTAGATCTATTATTTGTTCTCTTTGTGGGTAGGGTTT
GGTTAGTCCGTCCCGGATGACTATAAGCCGGCGTTTTTGTGCCCGGCGTCTCCGGTGGTCGGTCACTGGT
CAGTGACTCCGGCCGGTGAAGTCTCTGTGTCCATGGTTCTAGCTAGTTGCTGTTCCTTCCGCTGCTCTG

>**TaERF2-1** clone C9R1L1-5 1016bp CDS complete, mRNA 86.. 896 **B-1**

CAAAATCCACTGACAAACGCTCCGCGAATCCCAAGAGCGAACTCAGATCATCCTACGACCAGACGCGACC
ACACAGGATAAGATG**ATG**CTGCTTAATCCGGCGTCCGAGGCGGCGGCGCTGGACAGCATCCGGCAGCAGC
TCCTGGAGGAGCCGGCGCGGCCGGCGTACTGCCGGAGCGCGAGCTTCGGCAGCCTGGTGGCGGACCAGTG
GAGCGAGTCGCTCCCGTTCCGTCCCAACGACGCCGACGACATGGTCGTTTACGGCGCCCTCCGCGACGCC
TTCTCCTGCGGCTGGCTCCCCGACGGCCCCTTCGCGGCCGTCAAGCCCGAGCCCCTGCCCTCCCCCGACG
GCTCCTACGACGGCTCCTGCCTCGGCAGCTTCCTCGCGCCGCCGGCGCCCGGGCCGGACGCGCCGTGGGC

GGAGGAGGAGGCCGAGGTCGCGGCGGCGGCGTCGAGGGGGAAGCACTTCAGAGGCGTGAGGCAGCGGCCG
TGGGGCAAGTTCGCGGCGGAGATCCGGGACCCGGCCAAGAACGGCGCGCGCGTGTGGCTCGGCACCTTCG
ACAGCGCCGAGGACGCCGCCGTGGCCTACGACCGCGCCGCCTACCGCATGCGCGGCTCCCGCGCGCTCCT
CAACTTCCCGCTCCGCATCGGCTCCGAGATCGCCGCCGCCGCCGCAGCCGCGGGCCAGAAGCGTCCGTCT
CCCCAGCCGGCGAGCCCCGACTCTTCATCTCCCTCCTGCAGCGCGCCCGGGTCGTCGAAGAGGAGAAAGA
GAGGCGAGGCCGCGGCAGCGTCCATGGCCATGGCTCTGGTGCCGCCCCCGCCGGCGCAGGCTCCGGTCCA
GCTGACCCTCCCAGCCCAGCCGTGGCTGGCCGCCGGCGCCGTCCAGCAGCTGG**TGA**GCTGAAGCGGCGAA
GCGACCAGTGATCGTTCTCACTTCTCACGAGCGATTAGTTGCTTGATGTGTTGAGCGACGTGAGGAACAG
AGCATCAAGATGAGATCAATGGCGCCTAATGCTCGC

>**TaERF2-2** clone C9R1L1-6 1074bp CDS complete, mRNA 125.. 943 **B-1**

GAATTCCCGGGATATCGTCGACCCACGCGTCCGCGACTGGAGCACGAGGACACTGACATGGACTGAAGGA
GTAGAAAGAACTCAGATCATCTTTGGACCAGACGCGCCCACACAGGACAAGATG**ATG**CTGCTTAATCCGG
CGTCCGAGGCGGCGGCGGCGGCGGCGCTGGACAGCATCCGGCAGCAGCTCCTGGAGGAGCCGGCGCGGCC
GGCGTACTGCCGGAGCGCGAGCTTCGGCAGCCTGGTGGCGGACCAGTGGAGCGAGTCGCTCCCGTTCCGG
CCCAACGACGCCGACGACATGGTCGTCTACGGCGCCCTCCGCGACGCCTTCTCCTGCGGCTGGCTCCCCG
ACGGCTCCTTCGCAGCCGTCAAGCCCGAGCCCCTGCCCTCCCCCGGGTCCTACGACGGCTCCTGCCTCGG
CAGCTTCCTCGCGTCGCCGCCCGAGCTGGACGCGCCGTGGACGGAGGAGGAGGCCGAGGTCGCGGCGACG
GCGTCGAGGGGGAAGCACTTCAGAGGCGTGAGGCAGCGGCCGTGGGGCAAGTTCGCGGCGGAGATCCGGG
ACCCGGCCAAGAACGGCGCGCGCGTGTGGCTCGGCACGTTCGACAGCGCCGAGGACGCCGCGGTGGCCTA
CGACCGCGCCGCCTACCGCATGCGCGGCTCCCGCGCGCTCCTCAACTTCCCGCTCCGCATCGGCTCCGAG
ATCGCCGCCGCAGCCGCAGCCGCGGGCCAGAAGCGTCCGTCTCCGCAGCCGGCGAGCCCCGACTCTCCTC
CTCCCTCCTCAAGCGCACCCGGCTCGTCGAAGCGGAGAAAGAGAGGCGAGGCCGCGGCAGCGTCCATGGC
CATGGCTCTGGTGCCGCCCCCGCCGGTGCAGGCTCCGGTCCAGCTGACCCTCCCAGCCCAGCCGTGGCTC
GCCACCGGCGCCGTCCAGCAGCTAGTGAGC**TGA**AGCGGCGAAGCAACCACTGATCGTTCTCATGACCGAC
GGTTATTAGTTCTTCCTTCATGTGTTGAACCCACGGAGAAACAGAGCATCAAGATGAGATCAATGGCGCC
TAATGCTCGCAAGGGCGAATTGAT

>**TaERF2-3** clone C38F1L6-14 1166bp CDS complete, mRNA 216.. 812 **B-1**

ATCATTATTACGCCAAGCTCAGAATTAACCCTCACTAAAGGGACTAGTCCTGCAGGTTTAAACGAATTCG
CCCTTAACACCTCCCCTCCGACTCCCAGCAGCAGCCTTGCTCCCCTTGTTTCAACTACCCTTTGGATCCA
CCGATCGAGCGAGCGAGCCTTCTCCAGTTAGCCTTTCCCGCCACCTAAAAAAGCACCCGAGGTCGCCGTC
CATCC**ATG**TGCGGCGGCGCCATCATCTACGACTACATCCCGGCGCACCGCCGCCGGGTGTCCACCGCCGA
CTTCTGGCCCGACGCCAACGACCACTCCGACGCCCACAGCACCGCCCCCGACAAAGCGCCGCGCGCGAAA
CGGGGTGCGGACGAACCAGTACCGCGGCATCCGGCAGCGGCCGTGGGGGAAGTGGGCGGCGGAGATCCGC
GACCCCGTCAAGGGCGTCCGCGTCTGGCTCGGCACCTACCCCACCGCCGAGGCCGCCGCGCGCGCCTACG
ACCGCGCCGCCGCCGCATCAGGGGCGCCAAGGCCAAGGTCAACTTCCCCAACGAGATCCTCGTCGGCGCG
CCCGCGCACGAGGCCCCGTGCACGATTGCGGCCGTGCTCCCTTCCCCCAAGAAAGAGGAGGAGCCCGCGG
CGTGCTCCTGCGAGGAGGTGAAGGCGCTCTCCGAGGAGCTGATGGCGTACGAGAGCTACATGAGCTTCCT
CGGGGTCCCCTACATGGAGGGCGGGGCCGCGGCCGCCCCTGCCGCCGAGGAGGCGCCGGCCGAGCTATGG
AGCTTCGAGGACAGCTACTACCCAGGGCCTCTGGGGCTC**TGA**TTTTACCGTGCTCGTACTAGTCGCACCA
AACAGAGTCAACATTTTTCCCTTCTGTTCTGTTGAATTCATATTTTTTTTCAACTTGTAACTGTGGGATG
CATGCGCCAACGTGACTTTTACGATTTCTCCAAGTAAATAAAAAAAAAAAAAAAGGGGGGCGGCCCAGAGT
ATCCTCGAGGGGCCAAGCTACGCGACCACCTTCTTTAAAAATGTCCCATGTGAGCGTTTAAAACAGGCCT
GGGCCGTTACACGCGGACGGGAAACTCACTGGGTTAAGGGGAATTCGGCGCAAATAATTCCCTAAGGGCG
ATTAATTCGGCGCGTTTAAACCGGTGGGAAACCGGTACCACTATTC

>**TaERF2-4** clone WEF035_H08 977bp CDS complete, mRNA 205.. 786 **B-1**

CACCGGCGCCGCCTTCTCCCAGCTCCGACACAGCGGCAGCCCCCTCTGCGCAGAGTCAACAAGTCAAGCC
TTTCTCAAGCCCAGAGCAGGGTAGACGCCCGGCCTGACCAGCTCGGGGCAGCACCCTCGGCCTCTGGTGA
TAGTCACTAGCGAGGAAACAGCAGCCGGCGCGAGGAAGCGAGAGCCGGCCGGCGTTCTACTACT**ATG**AGG
AACGGCAGCACCGTGTGCGACGCGGCGGCGGCTGTGGGACTGGGCGTGGCGCCGCGGTACAGGGGCGTGC
GGAAGCGGCCGTGGGGCCGGTTCGCGGCGGAGATCCGGGACCCGGCCAAGCGGGCGCGGGTCTGGCTCGG
CACCTACGACTCCGCCGAGGCCGCCGCCCGCGCCTACGACGTCGCCGCGCGCACCCTCCGCGGCCCGCTC
GCCACCACCAACTTCCCCCACGACCTCGTCTCCTTCCCCACGCTCCCCGCGCTGCCGCGTCCGTCTCCAC

CGGCTTCCCCCGGCCCCGCCTGCAGCGCCAGCTCCACCGTCGAGTCCTACAGCGGGCCCAGGGCGGCGAA
CGCGCCTCGGGCCGTGCCGCGCGCGCCAGGGCCGACGCCCTTGCTGGCGAAGCTCCCAGGCGGCGACGCA
GGATGCCAGAGCGACTGCGCCTCCTCGGCCTCCGTCGTGGACGACGCCGGCGACGAGGCCGCCTCGGCCG
TCGTCCGGCCGCGCGCGCCCTTCGTGTTCAACCTCAACCTCCCGCCGCCGTCGGACGAGCTGTGCACGGA
GCTGCGGCTCTTC**TAG**GACGGATCCACCGGACCCGGATCGCCGGACGAGCTGTGCGCCAAAGCTAGCGAT
TTTAATTTTTTCTTCTATTCTTTTTCCTCCTGATTAGATTAGACAGATGATGTGTGCAAGAGAAGACCGA
TTGTACCACCAGCACTACTACTATTCTTTACGCGATAAGAATTGCGGTTGAAAAAAAAAAAAAAAAA

>**TaERF3-1** clone C65R1L1-83 902bp CDS complete, mRNA 84.. 788 **B-1**

ACTCCAAACCAAACCACCTGAACTCAACTCACCACTCAGCACAAGTCAGCGCCACCAAGCAAGAACACCA
CAAGAACAGACAC**ATG**ACCTTCAGCGTCTCGGCGGCGATGGAGGGCGGCCACGGCCACGAGTACATGATC
CGCTTCGACGGCCACTTGGAGGACCCGTCGCCCAGCACCGCCACTGCCGAGCCACCGCCGCCGCCGCCGC
CGTTCGCGGGCAGGGCGATCTCGCCGGAGCAGGAGCACGGCGCCATGGTCGCCGCGCTGCTGCACGTCAT
CTTCCGGGTACACCACGCCGCCGCCGGACTTCTTCCCGGCGGCCGCGGGCAAGGAGGTGTGCCCGGTGTG
CAGGGTCCACGGCTGCCTCGGCTGCGAGTTCTTCGGCGCGGCCGAGGCCACCGGGGCGACCGCGGCGGCA
TTTGGGACGCGCCGAAAGCAGCAATGACCGCGGGCGGGCCGCAGAGGCGGCGACGGAACAAGAAGAACAA
GTACAGGGGCGTCAGGCAGCGGCCGTGGGGCAAGTGGGCGGCGGAGATCCGCGACCCGCGCCGCGCCGTG
CGGGTGTGGCTCGGGACCTTCGACACCGCCGAGGACGCCGCCAGGGCCTACGACCGCGCCGCCGTCGAGT
TCCGCGGCCCGCGCGCCAAGCTCAACTTCCCCTTCCCCGAGCAGCAGCAGAACGGCGGGCAGGAAACAGG
GGATCAGCTCTGGGACGGCCTGCAGGACCTGATGAAGCTAGACGAGAGCGAGCTCTGGTTCCCGCCCTCT
GCAAATTCTTGGGAC**TGA**ACTGAACCTGCTTGATTAGATCCTAGCCGTTGGAGTGAGTGGACAAGACCAT
TTCACTTTTTTTTTCTTCCTTTTTTTTACCTCTGTTGCATTATTTGGACAAAACAGAGTCTG

>**TaERF3-2** clone WEF06_L20 817bp CDS complete, mRNA 142.. 762 **B-1**

CACCGGTCCGGAATTCCCGGGATATCGTCGACCCACGCGTCCGCCGGGCAGCTGCTTCCATCCCCCACCC
CAGATCCTGCACGCCATCCATCAGCTTGATACGCACACACCCATCCGTCAACAAGAAAAAGAAAGCCAGC
C**ATG**GCGCCCAGGACGTCCGACAAGACGGCGACGCCGCCCGCTGCCGGGGCCGCCGCGACCGGGCTCGCG
CTCGGCGTCGGCGGCGGCGGCAACGGCGGGGGCGTCGGGCCGCACTACCGCGGCGTCAGGAAGCGGCCGT
GGGGCCGGTACGCGGCGGAGATCCGCGACCCGGCCAAGAAGAGCCGGGTGTGGCTCGGCACGTACGACAC
CGCCGAGGAGGCCGCCCGCGCCTACGACGCCGCCGCCCGCGAGTACCGCGGCAACAAGGCCAAGACCAAC
TTCCCCTTCGCCTCCGCCTCCGCGCCGCCCGCCGCCGCGGCCCTCACCGGCGACGGCAGCCGGAGCAGCA
ACAGCAGCACCGTCGAGTCCTTCGACGGCGACGTGCAGGCGCCCATGCAGGCCATGCCGCTCCCTCCGTC
CCTCGAGCTCGACCTGTTCCACCGCGCGGCCACCAGCACCGCCGGGGCCGGCGCCGGCGCCGGCATGCGC
TTCCCCTTCAGCGGCTACCCCGTGTCGCACCCGTACTACTTCTTCGGACAGGCGGCGGCCGCTCTAGAGT
ATCCCTCGAGGGGCCCAAGCTTACGCGTACCCAGCTTTCTTGTACAAAGTGGTCCCCTA**TAG**TGAGTTCG
TATTATAAGCTAGGCACTGGCCGTCGTTTTAAAACGTCGTGTACTGG

>**TaERF3-3** clone CA652238R1L1-57 960bp CDS complete, mRNA 78.. 800 **B-1**

CCACCCCAGATCCTGCACGCCATCCATCAGCTTGATACGCACACACCCATCCGTCAACAAGAAAAAGAAA
GCCAGCC**ATG**GCGCCCAGGACGTCCGACAAGACGGCGACGCCGCCCGCTGCCGGGGCCGCCGCGACCGGG
CTCGCGCTCGGCGTCGGCGGCGGCGGCAACGGCGGGGGCGTCGGGCCGCACTACCGCGGCGTCAGGAAGC
GGCCGTGGGGCCGGTACGCGGCGGAGATCCGCGACCCGGCCAAGAAGAGCCGGGTGTGGCTCGGCACGTA
CGACACCGCCGAGGAGACCGCCCGCGCCTACGACGCCGCCGCCCGCGAGTACCGCGGCAACAAGGCCAAG
ACCAACTTCCCCTTCGCCTCCGCCTCCGCGCCGCCCGCCGCCGCGGCCCTCACCGGCGACGGCAGCCGGA
GCAGCAACAGCAGCACCGTCGAGTCCTTCGACGGCGACGTGCAGGCGCCCATGCAGGCCATGCCGCTCCC
TCCGTCCCTCGAGCTCGACCTGTTCCACCGCGCGGCCACCAGCACCGCCGGGGCCGGCGCCGGCGCCGGC
ATGCGCTTCCCCTTCAGCGGCTACCCCGTGTCGCACCCGTACTACTTCTTCGGACAGGCGGCGGCCGCCG
CCGCCGGCGGCTGCCACATGTACAGCCAGGCCCCGAAGGTGACCGTGGCGTCCGTGTCCCCGAGCGACTC
CGACTCCTCGTCGGTGGTGGATCTGGCGCCGTCGCCGCCCTCAAGGAAGCCCGTCCCTTTCGACCTCGAC
CTGAACTGCCCGCCGCCGGCGGAGCTC**TGA**TCGGGCGCCGGAGTTTATGTAGCTGATATGACTGCTAGTT
TCTTTCGTCGCCTTCTTTTTTTGCCCCTAGTAGAACGAAGAAAAATTTGCATGTACCTCCATGATGTTTT
TAGAAGAGGCCGGCCGCCTTGCACAACAACGGCAGGGGCAGTTCTGTAAA

>**TaERF3-4** clone bq48379f1l1-31.g 767bp CDS complete, mRNA 9.. 668 **B-1**

CCGCTCCC**ATG**GCTCCCAAGAACGCGCTCCCCGTCGCCGTCGCCGCCGCCGCAGACGGCGGCATGGAGCC
CAGGTTCCGCGGCGTGCGGAAGCGGCCGTGGGGCAGGTACGCGGCGGAGATCCGCGACCCGGCCAGGAAG
GCGCGCGTCTGGCTCGGCACCTTCGACACCGCCGAGGCCGCCGCGCGCGCCTACGACGCCGCCGCGCTCC
ACTACCGCGGGCCCAAGGCCAAGACCAACTTCCCCGTCGGCACCGTCGCCGCCTACACGCACATCCCGCT
CCCGCCGCCCAAGGCGCTGGCCGTGAGCCCCAGCAGCAGCACCGTCGAGTCGTCGTCCCGGGACACGCCG
GCCGCGGCGCCCGCCGCTCCTGCTGTTGCTGCCCCCGCGGCCCCGCCGGCGCTCGACCTGAGCCTGGCGA
TGCCGGCCATGGTGGCGGCGCAGCCGTTCCTGTTCCTGGACCCCAGGGTCGCGGTGACCGTGGCCGTGGC
AGCGCCGGCGCCGGCGCCGGCGCCGTGCCGGTCAGCGGCGATCAGCGGCATGAACAAGGTGGCGTCCCAC
GAGGAAGAGCAGAGCGACACCGGGTCGTCGTCATCCGTGGTGGACGCCTCGCCGGCCGTGGGCGTGGGGT
TTGACCTGAACCTGCCGCCGCCgGTGGAGATGGCA**TAG**GAGATGGACCGATCTCGGTCGCCCGATGACgA
CgCGGCCCGGAAGTaAAGCAAGCTCTCTCGTGGTAgATTTTAAAGTTTAAAAgATcGGCGTGTACTA

>**TaERF3-5** clone CA652238R1L1-54 979bp CDS complete, mRNA 113.. 823 **B-1**

GTCCCTCCTCCCCATCCCTCTCCCACCCCCCCACCCCAGATCCTGCACGCCATCCATCAGCTTCATATGC
ACACACCCATCCGTCAACAAGAAAAAGAAAGCCAAGCCAGCC**ATG**GCGCCCAGGACGTCCGACAAGACGG
CGACGCCGCCCGCTGCCGGGGCCGCCGCGACCGGGCTCGCGCTCGGCGTCGGCGGCGGCAACGGCGGGGG
CGTCGGCCCGCACTACCGCGGCGTGAGGAAGCGGCCGTGGGGCCGGTACGCGGCGGAGATCCGCGACCCG
GCCAAGAAGAGCCGGGTGTGGCTCGGCACGTACGACACGGCCGAGGAGGCCGCCCGGGCCTACGACGCCG
CCGCGCGCGAGTACCGCGGCAACAAGGCCAAGACCAACTTCCCCTTCGCCACCGCCTCCGCGCCGCCCGC
CGCCGCGGCGGCCCTCACCGTCGACGGCAGCCGGAGCAGCAACAGCAGCACCGTGGAGTCCTTCGGCGGC
GACGTGCAGGCGCCCATGCAGGCCATGCCGCTCCCTCCGTCCCTCGACCTCGACCTGTTCCACCGCGCGG
CCACCAGCACCGCCGGCGCCGGCATGCGGTTCCCCTTCAGCGGCTACCCCGTGTCGCACCCGTACTACTT
CTTCGGACAGGCCGCGGCCGCGGCCGCCGCCGGCTGCCACATGTACAGCCAGGCCCCGAAGGTGACCGTG
GCGTCCGTGTCCCCGAGCGACTCCGACTCGTCGTCGGTGGTGGATCTGGCGCCGTCGCCGCCCGCGAGGA
AGCCCGTCCCTTTCGAGCTCGACCTGAACTGCCCGCCGCCGGCGGAGCTC**TGA**TCGGGGCCGCCGGAGTT
TATGTAGCTGATATGACTGCTAGTTTCTTTCGTCGCCTTCTTTTTTTGCCCCTAGTAGGAAGGAAATTTC
CATGTACTCCATGATGTTTAGGAGAGGCCGTCCGCCTTGCACAACAACGGCAGGGGCAGTTCTGTAAAT

>**TaERF3-6** clone c37fll8-d 1268bp CDS complete, mRNA 88.. 798 **B-1**

TTCGCCCTTGATCCTGCACGCCATCCATCAGCTTCATACGCACACACCCATCCGTCAACAAGAAAGAGAA
AGGCGAGCCGGCCAGCC**ATG**GCGCCCAGGACGTCCGACAAGACGGCGACGCCGCCCGCTGCCGGGGCCGC
CGCGACCGGGCTCGCGCTCGGAGTCGGCGGCGGCAACGGCGGGGGCGTCGGCACGCACTACCGCGGCGTC
AGGAAGCGGCCGTGGGGCCGGTACGCGGCGGAGATCCGTGACCCGGCCAAGAAGAGCCGGGTGTGGCTCG
GCACGTACGACACGGCCGAGGAGGCCGCCCGCGCCTACGACGCCGCCGCGCGCGAGTACCGCGGCAACAA
GGCCAAGACCAACTTCCCCTTCGCCTCCGCCTCCGCGCCGCCCGCCGCCGCCGCGGCCCTCACCGGCGAC
GGCAGCCGGAGCAGCAACAGCAGCACCGTCGAGTCCTTCGGCGGCGACGTGCAGGCGCCCATGCAGGCCA
TGCCGCTCCCTCCGTCCCTCGAGCTCGACCTGTTCCACCGCGCGGCCACCAGCACCGCCGGCGCCGGCAT
GCGCTTCCCCTTCAGCGGCTACCCGGTGTCGCACCCGTACTACTTCTTCGGGCAGGCCGCGACCGCCGCC
GCCGCCGGCTGCCACATGTACAGCCAGGCGCCGAAGGTGACCGTGGCGTCCGTGTCCCCGAGCGACTCCG
ACTCCTCGTCGGTGGTGGATCTGGCGCCGTCGCCGCCCACAAGGAAGCCCGTCCCGTTCGACCTCGACCT
GAACTGCCCGCCGCCGGCGGAGCTC**TGA**TCGGGGCCGCCGGAGTTCATGTAGCTGATATGACTGCTAGTT
TCTTTGGTCGCCTTCTTTTTTTGCCCCTAGTAGAGCGAAGAAAATTTGCATGTACCTCCATGATGTTTTT
AGGAGAGGCCGTCCGCCTTGCACAACAGCGGCAGGGGCAGATCTGTAAATACGGTTTCTTTTTTCGCCGG
AGAGGCAGAGATGGATGGATGTATTGTCGCCCAAGACAAGTCCTGCCGGCGTTGCCATAGACCGAATTAT
ATCGAGTTACTTCCTACTTTGACAAAAAAAAAAAAAGGGCGGCCGCTCTAGAGTATCCCTCGAGGGGCCCA
AGCTTACGCGTACCCAGCTTTCTTGTACAAAGTGGTCCCTATAGTGAGTCGTATTATAAGCTAGGCACTG
GCCGTCGTTTTACAACGTCGTGACTGGGAAAACTGCTAGCTTGGGATCTAAGGGCGAATTCGTTAAACTG
CAGACTCC

>**TaERF3-7** clone C68R1L1-7 1055bp CDS complete, mRNA 97.. 843 **B-1**

GCGCTACACACCCACCCCACCCCACCCCCCTGCATCACCAAAACGCCCGCATAACCAAAAAGATACGCGC
ACGCCACTTGTACACCTGCTGCACCC**ATG**GCGCCTAGAGCGGCGGAGAAGGCGCCTGTCTCCCCGCCCAC
CGGCCTCGGCCTCGACGTCGGCGGCGGGATCGGGGTCGTCGCCGGCGGCGCGCACTACAGGGGCGTCCGG
AAGCGCCCGTGGGGACACTACGCCGCGGAGATCCGCGACCCGGCCAAGAAGAGCAGGGTGTGGCTCGGCA
CGTACGACACCGCGGAGGAGGCCGCGCGCGCCTACGACACCGCCGCCCGCGAGTTCCGCGGCGCCAAGGC

GAAGACCAACTTTCCCGTTCCCTTCGTCGTCGTTGTCCCCCTGTCGCCGCCGCCGCCGGCGGCGGCGGCA
GCCCCGAGCAGCGACAGCACCTTGGACTCGAGCGGCGGAGGGAGCGGCGGCTGCGCCCAGGCGCCCATGC
AGGCCATCCCGCTGCCGCCCGCCCTCGACCTGGACCTCTTCCACCGCGCGGCGGTCGTCACGGCCGTCGC
CGGCGGCGGCATGCGCTTCCCATTCAACGGTTACCCGGTGGCGCCGCGCCAGCCCCTGCACCCGTACTTC
TTCTACGAACAGGCCGCGGCCGCCGCGGCGGCGTCGGCGTCGGGGTACCGCGCGCTGAAGGTGGCGCAGC
CGGTCACCGTGGCGGCCGTTGCGCAGAGCGACTCCGACTCCTCGTCGGTCGTTGATCTGTCCCCGTCGCC
CCCCGCGGTGACAGCGCATAAGGCGGTCGCGTTTGATCTGGATCTCAACCGGCCGCCGCCTTCCGAGGAC
TAGACAAAGGACAAATTTTAGATGATGACTGTAGTTTCTCATTTTTCCTGCAGGGAACGCTTTTTTTCCT
TCTTCTGTTCTGGTCCTCTTGTATTTTTGTTTAGTTGTCCGAGACAGACCGAGGAGCCCTGTAAATAGTT
TTTCCGCCGAGAACAGAGCAGAACCGATCTGAGATCTGTTGGTCTAAAACGGATCAAACCGGCGCTGAGG
GAACC

>**TaERF3-8** clone BQ483794F1L1-31 698bp CDS complete, mRNA 9.. 668 **B-1**

CCGCTCCC**ATG**GCTCCCAAGAACGCGCTCCCCGTCGCCGTCGCCGCCGCCGCAGACGGCGGCATGGAGCC
CAGGTTCCGCGGCGTGCGGAAGCGGCCGTGGGGCAGGTACGCGGCGGAGATCCGCGACCCGGCCAGGAAG
GCGCGCGTCTGGCTCGGCACCTTCGACACCGCCGAGGCCGCCGCGCGCGCCTACGACGCCGCCGCGCTCC
ACTACCGCGGGCCCAAGGCCAAGACCAACTTCCCCGTCGGCACCGTCGCCGCCTACACGCACATCCCGCT
CCCGCCGCCCAAGGCGCTGGCCGTGAGCCCCAGCAGCAGCACCGTCGAGTCGTCGTCCCGGGACACGCCG
GCCGCGGCGCCCGCCGCTCCTGCTGTTGCTGCCCCCGCGGCCCCGCCGGCGCTCGACCTGAGCCTGGCGA
TGCCGGCCATGGTGGCGGCGCAGCCGTTCCTGTTCCTGGACCCCAGGGTCGCGGTGACCGTGGCCGTGGC
AGCGCCGGCGCCGGCGCCGGCGCCGTGCCGGTCAGCGGCGATCAGCGGCATGAACAAGGTGGCGTCCCAC
GAGGAAGAGCAGAGCGACACCGGGTCGTCGTCATCCGTGGTGGACGCCTCGCCGGCCGTGGGCGTGGGGT
TTGACCTGAACCTGCCGCCGCCGGTGGAGATGGCA**TAG**GAGATGGACCGATCTCGGTCGCCCGATGAC

>**TaERF3-9** clone CA683704F1L8-8 1108bp CDS complete, mRNA 68..797 **B-1**

CTGCATCACCAAAACGCCCGAATAACCAAAAAGATACGCGCACGGCACTTGTACACCTGCTGCACCC**ATG**
GCGCCTAGAGCGGCGGAGAAGGCGCCTGTCCCCCCGCCCACCAGGCTCGGCCTTGGCGTTGGCGGCGGGG
TCGGAGTCGTCGCCGGTGGCGCGCACTACAGGGGCGTCCGGAAGCGCCCATGGGGACGTTACGCCGCGGA
GATCCGTGACCCGGCCAAGAAGAGCAGGGTCTGGCTCGGCACGTACGATACGGCGGAGGAGGCCGCGCGC
GCCTACGACACCGCCGCGCGCGAGTTCCGCGGCGCCAAGGCGAAAACTAACTTCCCGTTCCGTTCGTCGT
CGTCGTCGTCCCCTGTCGCCGGCGGGGGCAGCCCGAGCAGCAACAGCACCTTGGACTCTAGCGGGGGAGG
GAGCGGCGGCTGCGCCCAGGCGCCCATGCAGGCCATCCCGCTGCCGCCCGCCCTCGACCTGGACCTCTTC
CACCGCGCGGCGGGCCGTGACGGCCGTCGCCGGCGGCGGCATGCGCTTCCCGGTCAACGGTTACCCGGTG
GCATCGCGCCAGCCTCTGCACCCGTACTTCTTCTACGAACAGGCCGCGGCCGCCGCGCGGCGGCGTCGNG
GGTATCGCGCGCTGAAGGTGGCGCACGGTTCACCGTGGCGGCCGTTTGCCAAAGCTACTCCGACTCCTCG
CCGGTCGTTTATTTTTCCCCTTCGCCCCCCGGGTGACATCGCATAAAGGCGGCACGTTTGATTTGGATCT
TAACGGGCCCGCCCTTCGGAGGAC**TAG**AAAAAAGAGAAAATTTTGAATGTGACTGTGTGTGTTCATTTTT
CTGTGAGGAATTTTTTTTTCTTCCTCTCGGTCGGGGCCCCTGGATTTTGTTTGAGTTTTCGCAAAAACGG
GAGCCTGTAAAAATTTTTCGCGCGAAACGACAAACCGGTGAGTTTGTTGTCTAAACGACAAACCGCGCTC
ACGGGACCGCACTTTGTTTATCGTGTAAACACTGTGTCTGATGTGTTTCTAAAAAGTGTCTCTCTTGCAA
AAAAAAAAAGGGGGCGCTCGATATCCCCAGGGGGAGTGTGTGACCCTTTTTTAAAATG

>**TaERF7-1** clone WEF051_E24 824bp CDS partiale, mRNA 178.. >824**B-1**

ACGCTTGACATTAGGCCTATTTAGGTGAACTATAGAACAAGTTTGTACAAAAAAGCAGGCTGGTACCGGT
CCGGAATTCCCGGGATATCGTCGACCCACGCGTCCGAGACCATCCAACCAACGCCGCACCGCAACAACAA
CCGCCGCGCACGCGCCGAGCACGGTCGAACACTTGCC**ATG**GCTCCCAAGAACGCGCTCCCCGTCGCCGTC
GCCGCCGCCGCGGACGCCGGCATGGAGCCCAGGTTCCGCGGCGTGCGCAAGCGGCCGTGGGGCAGGTACG
CGGCGGAGATCCGCGACCGNGCCAGGAAGGCCCGCGTCTGGCTCGGCACCTTCGACACCGCCGAGGCCGC
GGCCCGCGCCTACGACGCCGCCGCGCTCCACTACCGCGGGCCCAAGGCCAAGACCAACTTCCCCGTCGGC
ACCGTCGCCGCCTTCGCCCACGTCCCGCTCCCGCCGCCCAAGGCGCTGGCCGTCAGCCCCAGCAGCAGCA
CCGTCGAGTCTTCGTCCCGGGACACGCCGGCCGCGGCGCCCGCCGCTCCTGCTGCTGCTGCCCCCGCGCC
CCCGCCGGCGCTCGACCTGAGCCTGGCGATGCCGGCCATGGTGGCGGCGCAGCCGTTCCTGTTCCTGGAC
CCCAGGGTCGCGGTGACCGTGGCCGTGGCCGCGCCGGCGCCGGCCCCCTGCCGGTCAGCGGGGATCAGCG
GGATGAACAAGGTGGGCGTCCGCGAAGAAAAGCAAAGCGACACCGGGTCGTCCTCATCCGGGGGTGGAAC
CCCTCCCCGCCGGGGGGCGGGGGGTTCAACCCTGAACCTGCCCCCGCCCGGGGA

>**TaERF7-2** clone WEF53A07 958bp CDS complete, mRNA 178...765 **B-1**

CAGCTCCAGCACCGCGGCTCGGCTCCCTCTACCCAGAGCCAAAGAATCAAGGCTGTTTGAAGCTCGAGAT
CACGCGCGTTTCATTCAGCCCCTCTGGTGGTGGTAGCCAGGCCGAGACAGTCACCGGGCGGAAGGAGCGG
AAGGACTGAAGCCTCAGCGGGCCGGCGATCTACTACT**ATG**AGGAACGGCAGCACCGTGTGCGACGCGGCG
GCGGCTGTGGGGCTGGGCGTGGCGCCGCGGTACAGGGGGGTGCGGAAGCGGCCGTGGGGCAGGTTCGCGG
CGGAGATCCGGGACCCGGCCAAAGCGGGCGCGGGTCTGGCTCGGCACCTACGACTCCGCCGAGGCCGCCG
CGCGGGCCTACGACGTCGCCGCGCGCACCCTCCGCGGCCCGCTCGCCACCACCAACTTTCCCCAGGACCC
CGTCTCCTTCCCCACGCTCCCCACGCTCCCGCATCCGTCTCGACCGGCTTCCCCCGGCCCAGCCTGCAAG
CGCCAGCTCCACCGATCGAGTCCTTACAGCGGGCCCAGAGCCTGCGAAATGCGCCTCGGGCCGTCCCGCG
CGTGTCAGCGCCGGCACCTTTGCTGGCGGAGCGCTCCGGCGGCGACGCAGGATGCCAGAGCGACTGCGCC
TCGTCGGCCTCCGTCGTGGACGACGGCGGCGACGAGGCCGCCTCGGCCGCCGTCCGGCCTCGCGCGCCCT
TTGTGTTCGACCTCAACCTCCCGCCGCCGTCGGACGAGCTGTGCACCGAGCTACGGCTTTTC**TAG**GATAC
ACCGGACCCGGGTCGCCGGACGAGCTGCGCACCAAAGCTAGCGATTTTTTATTTTTCTTCTATTCTTTTT
CCTCCTGATTATTAGATTAGACAGATGATGCTGTGTGCAAGAGCGATTGTACCACCACCACTACTATTCT
TTACGCGATAAGAATTGCGGTTGATTTGATTGAAAAAAAAAAAAAAAA

>**TaERF7-3** clone WEF86p16 1180bp CDS complete, mRNA 227...860 **B-1**

GAAGACACCGGGACCGATCCAGCCTCCGGACTCTAGCCTAGGCCGCGGGACGGATAACAATTTCACACAG
GAAACAGCTATGACCATTAGGCCTATTTAGGTGACACTATAGAACAAGTTTGTACAAAAAAGCAGGCTGG
TACCGGTCCGGAATTCCCGGGATATCGTCGACCCACGCGTCCGCGCCCAGCCACCGCCACAACAACCGCC
GCGTCGAACACTTGCC**ATG**GCTCCCAAGAACGCGGCCCCCGCCGCCGCCGCCGCGGACGCCGGCATGGAG
CCGAGGTTCCGCGGCGTGCGGAAGCGGCCGTGGGGCAGGTACGCGGCGGAGATCCGCGACCCGGCCAGGA
AGGCGCGCGTCTGGCTCGGCACCTTCGACACCGCCGAGGCCGCCGCGCGCGCCTACGACGCCGCCGCGCT
CCACTACCGCGGGCCCAAGGCCAAGACCAACTTCCCCGTCGGCACCGTCGCCGCCTTCGCGCACGTCCCG
TTCCCGCCGCCCAAGGCGCTGGCCGTCAGCCCCAGCAGCAGCACCGTCGAGTCCTCGTCCCGGGACACGC
CGGCCGCGCCGCCCGCCGCCCCGCGCCCCCGCCGGCGCTCGACCTGAGCCTGGCGATGCCGGCCATGGT
GGCGGCGCAGCCGTTCCTGTTCCTGGACCCGAGGGTCGCGGTGACCGTGGCCGTGGCCGCGCCGGCGCCG
GCGCCGTGCCGGTCGGCGGCGATCAGCGGCATGAACAAGGCGGCGTCCCGCGAGGAGGAGCAGAGCGACA
CGGGGTCGTCGTCATCCGTGGTGGACGCCTCGCCGGCCGTGGGCGTGGGGTTCGACCTGAACCTGCCGCC
GCCGGTGGAGATGGCG**TAG**GAGATGGACCGATCTCGGTCGCCCGATGACGACGCGGCCCGGAAGCAAAGC
GAAGCTCTCTCTTGGGTAGAACTTAAAGTTTAGAAGAATGGCGTGTACTTACTAGTAAGAGGTTAATCGG
CCGTTTAATTGCCAGGATCGATCCATCGATGGGTAAGGGTAAGTAGGTGTACATAGATCGCGTGCGTGCG
TGCGTTTTTTGCATGGGGAGGCAGGTAGCATGATGAGGAACATCCATCTATCTGCTCCCGATGTTACTTA
CAATTTCCTTAAAAGCAAGGGAAAAACAAAGACAAAAATGGAAGAAAAAAAAAAAAAAAA

>**TaERF-1** clone C72R1L8-2 1061bp CDS complete, mRNA 32.. 994 **B-1** rootAP

AAAACCCAACGGAAAAAGCAAAGAGAGAACG**ATG**TGCGGCGGCGCCATCCTAGCGCAGCTGATCCCGCCG
TCGGCGGGCCGTGCCCCGAAGCAGGTGGCCGCCGGCGGGGTCGCGCCCAAGAATGGCGGCATGAGCAAGA
GGCACCACAGCAGCACCCCCGATGTCGACGACTTCGAGGCCGCCTTCGAGGACTTCGAGGACGACGTCGA
CCTGCAGGCGGAGGACGACGGCGACGACCATGCCGTTTTTGCATCCAAGCCCGCCTTCTTCCACGGGCCT
AATTATGATGGCCGCGCGGCGCAGGCGGCCAGCAGGAAGAAGAGCGTCCGCCGCCTCCCACGGCATCCGG
CAGCGGCCGTGGGGCAAGTGGGGCGGCGGAGATCCGCGACCCCGCACAAGGGCACCCGCGTCTGGCTCGG
AACGTTCGACACGGCCGATGACGCCGCCCAGGCTTACGACGTCGCCGCCCGCCGCCTCCGTGGCATCAAG
GCCAAGGTCAACTTCCCCGACGCGGCGAGGGCCGGGGCGCGCTCGCGCCGCGCCAGCCGGAGAACTGCGC
AGAAACCGCAGTGCCCGCCTGTGCGGACGACGGCGTACTCTGCCACCGCAGCTGCACACGCACACGTAGT
ACAGGCGGAGCAGGACGCTATGATGATCAAACCTGAGCTGATGGAGTTTTTTGATGCGGACGCGTTCGTC
GACCTGACCGCCGCCGTCGCCGCGCTGCCGCCTGTCACTGGCGCGAAGAAGCCGATGGTCGATGAGGATT
CGTCGGATCGGAGCGGCGGCTGCGCCATGCTGGTGTTCGCCGACGAGCTTGGGTTCGATCCGTTCACGCT
GTTCCAGCTCCCCTGCTCGGACACCTACGAATCCATCGACAGCCTCTTCGCCGGGGACGCCGTCATCCAG
GATGCCCTCGGCGTGGACACTGGCATGGAGGGCGTCAGCCTCTGGAGCTTCGAGGAGTTCCCCATGGACA
GCGCCATTTTT**TGA**CGTTTCCGTGCGATGCGCCCGATCGGTTGTAAGAATCTCCATCTGGCATCTTGTTC
CATGGTGCACA

>**TaERF-2** clone C82R1L8-8 1174bp CDS complete, mRNA 28.. 978 **B-1** rootAP

AAAAGCGCGAGACAGAAGGAGAAAACA**ATG**TGCGGCGGCGCCATCCTAGCGGAGCTGATCCCGCCGTCGG
CGGGCCGTGCCTCGAAGCAGGTGGCCGCGGGCCGGGCCTCGCCAAAGAAGGCCGGCAAGAGCAAGGGGCA
TAGGTACGGCAGCGTCGCCGATGTCGACGACTTCGAGGCCGCCTTCGAGAACTTCGACGACGACCTCGAC
CTGCAGGCGGAGGAGGACGGCGACGAACATGTCGTTTTTGCATCCAAGCCTGCGTTCTCTCCGGGCCTGG
CCTACGATGATGGCCGCGCGGCGCAGGCGGCGAGCAAGAAGAAGAGCGTCCGCCCCCTCCACGGCATCCG
GCAGCGGCCGTGGGGTAAGTGGGCGGCGGAGATCCGCGACCCGCACAAGGGCACCCGCGTCTGGCTCGGC
ACCTTCGACACGGCCGACGATGCCGCCCGGGCCTACGACGTCGCCGCCCGACGCCTCCGCGGCAGCAAGG
CCAAGGTCAACTTCCCCGACGCGGCCAGGGCCGGGGCGCGCCCGCGCCGCGCCAGCCGCAGAACTGCGCA
GAAACCGCAGCGCCCGCCTGCGCGGACGACGGCGTACTCTGCCACCGCAGCACCACGCGCACGGCCGGAG
CAGGACGCTATGATGGTCAAGCCCGAGCGGATGGAGTTTTCCGACGTGGACGCGTTCGTCGACCTCACCG
CCGCCGTCGCCGTGCTACCGCCTGTCACGGCGAGCTCCTTCGCCGACAAGATGCCGAGGGTCGACGAGGA
CTCGTCGGAGGGGAGCGGCGGCGGCTCCATGTTGGGGTTCGCCGACGACCTTGGGTTCGATCCCTTCATG
ATGCTCCAGTACGAATCCATGGACAGCCTCTTCGCCGGAGACGCCATCATCCAGGATGCCCGCGGTGTGG
ACGGCGGCATGGACGGCGTCAGCCTCTGGAGCTTCGAGGAGTTCCCCATGGACAGCGCCATTTTC**TGA**CG
TTTTCCGTGCGATGCGCTCGATCGGTTGAACGAAGGAAAGGGTGCTGCCGATGCATGGCTACCGCATGAT
TGCTTCCTTGATGAACGCAGATTCGAAATGTATTCCACTGTTTGAGGCTGCTCGTCGGACTGTATGTGCA
TATGTACTGCTGGCCGCCTGTACTCTTTCATGATTTGTGCTAGCCTTTGGGTGC

>**TaERF-3** clone c86r1L1-76 1283bp CDS complete, mRNA 118. 1068 **B-1** rootAP

GGAATTCCCGGGATATCGTCGACCCACGCGTCCGCGACTGGAGCACGAGGACACTGACATGGACTGAAGG
AGTAGAAAGTGAAAGAAAAATCCAACGGAAAAGGCAAAGAGAGAACA**ATG**TGCGGCGGCGCCATCCTAGC
GAAGCTGATCCCGCCGACGCCGCCGTCGGCGGGCCGTGCCCCGAAGCAGGTGGCCGCGGGCGGGGTCTCG
CCCAAGAAGGGCGGCATGAACAAGACGCACCACAGCAGCACCCCCGATGTCGACGACTTCGAGGCCGCCT
TCGAGGACTTCGATGACGACTTCCACCTGCAGGCGGAGGAGGACGGCGACGACCATGTCGTTTTTGCATC
CAAACCTGCCTTCTCCCCGGCCTACGATGATGGCCGCGCGGCGCAGGCGGCGAGCAGGAAGAAGAGCGTC
CGCCGCCTCCACGGCATCCGTCAGCGGCCGTGGGGCAAGTGGGCGGCGGAGATCCGCGACCCGCACAAGG
GCACCCGCGTCTGGCTCGGCACGTTCGACACGGCCGATGATGCCGCCCGGGCCTACGACGTCGCCGCCCA
CCGCCTCCGTGGCAGCAAGGCCAAAGTCAACTTCCCCAACGGGACCAGGGCTGGGGCGCGCCTGCAACGT
GCCAGCCGGAGAACCGCTTCGAAACGGCAGTGCCCCCCTGCGCGGACGACGGCGTACTCTGCTGCACACG
CACAGAAGGAGCGGGACGCTATGGTGGCCAAGCCTGAGCTGATGGAGTCTTTCGACATGGACGCCTTCGT
CGACCTGACCACTGCTTTCACCACGCTACCGCCTGTCATGGCAAGCTCCTTCGCCGACACTGGCGCGAAG
AAGCCGATGGTCGATGAGGATTCGTCGGATGGGAGCGGCGGCGATGCCATGCTGGGGTTCGATCCGTTCA
TGCTGTTCCAGCTCCCCTGCTCGGACACGTACGAATCCATCGACAGCCTCTTCGCCGGCGACGCCGTCAT
CCAGGATGCCCTCGGCGTGGACAGTGGCATGGAGGGCGTCAGCCTCTGGAGCTTCGAGGAGTTCCCCATG
GACAGCGCCATTTTT**TGA**CGTTTCCGTGCGATGCGCCCGATCGGTTGAACGAAGGGCGCAGCCGATGCAT
GGCTGCTGCATGATTGCTTCCTTGATGAATGCAGATTCGTCGGACTGTATGTGCATATGTACTGCTGGCC
GCCTGTATCAAACTGTATTGTACTCTGCCATGATTTGTGCTAGTCTTTCCTACTGTTGGTCAATAGTCCT
ACGAATTGATCATGTCGGCAGTG

>**TaERF-4** clone C87R1L6-67 1370bp CDS complete, mRNA 29.. 976 **B-1** rootAP

AAAAAGCGCCAGACAGAAAGAGAAAACA**ATG**TGCGGCGGCGCCATCCTAGCGGAGCTGATCCCGCCGTCG
GCGGGCCGTGCCTCGAAGCAGGTGGCCGCGGGCCGGGCCTCGCCAAAGAAGGCCGGCAAGAGCAAGGGGC
AGAAGTACGGCAGCGTCGCCGATGTCGACGACTTCGAGGCCGCCTTCGAGAACTTCGATGACGACCTAGA
CCTGCAGGCGGAGGAGGACGGCGACGACCATGTCGTTTTCGCATCCAGGCCTGCGTTCTCCCCGGCCTAC
GATGGTGGCCGCCGCGCGGTGCCGGCGGCGAGCAAGAAGAAGAGCGTCCGCCCCCTCCACGGCATCCGGC
AGCGGCCGTGGGGCAAGTGGGCGGCGGAGATCCGCGACCCGCACAAGGGCACCCGCGCCTGGCTCGGCAC
CTTCGACACGGCCGATGATGCCGCCCGGGCCTACGACGTCGCCGCCCGTCGCCTCCGTGGCAGCAAGGCC
AAGGTCAACTTCCCCGACGCGGCCAGGGCCGGGGCGCGCCCGCGCCGCGCCAGCCGTAGAACCGCGCAGA
AACCGCCATGCCCCCCTGCGAGGAGGACGGCGTACTCTGCCACCGCAGCAGCACGCGCACAGCCGGAGCA
GGACTCTATGATGGTCAAGCCCGAGCTGATGGAGTCTTTAGACATGGACGCCCTTGTCGACCTGACCACT
GCTGTCACCGCACTACCGACTGTCATGGCAAGCTCCTTCGCCGACAAGATGCCGAGGGTCGACGAGGACT
CGTCGGAGGGGAGCGGTGGCGGCGCCATGCTGGGGTTCGCCGACGAGCTTGGGTTCGATCCGTTCATGAT
GTTCCCGTACGAATCCATGGACAGCCTCTTCGCCGGAGACGCTGTCATCCAGGATGCCCGCGGTGTGGAC

GGCGGCATGGACGGCGTTAGCCTCTGGAGCTTCGACGAGTTCCCCATGGACAGCGCCATTTTC**TGA**CGTT
TTCCGTGCGATGCGCTCGATCGGTCGAACCAAGGAAGGGTGCAGCTGATGCATGGCTACTGCATGATTGC
TTCCTTGATGAATGCAGATTCGAAATGTATCCCACTGTTTGAGGCTGTTCGTCGGACTGTATGTGCATAT
GTACTGCTGGCCGCCTGTACTATCAAGCTGTATTGTACTCTGTCATGATTTGGTACTGTTGGTCAATAGT
CCTACGAATTGATCATGTCGGCAGTGCCGTCTATGATTGTGCTTGCCATTGTTAATGCCAAATAAGCTCG
TGAACCTGGTTATCTGTTGTGTGTGAGTCACCAACTCATCATGTCATGTTGAGTTGGATTATGTAATAAA
TACTGTAATTTGCTGAATTCGGTGAAGCTGTACAATCCTG

>**TaERF-5** clone C88R1L1-79 1201bp CDS complete, mRNA 13.. 954 **B-1** rootAP

AAAGAGAAAACA**ATG**TGCGGCGGCGCCATCCTAGCGGAGCTGATCCCGCCGTCGGCGGCGGGCCGTGCGC
CGAAGCAGGTGGCCGCGGGCCAGGCCTCGTCCAAGAAAGGCGGCATGAGCAAGAGGCACCACAGCAGCAT
CCCCGACGTCGACGACTTCGAGGCCGCCTTCGAGGACTTCGATGACGACTTCGACCTGCAGGCGGAGGAG
GACGGCGGCGACCATGTCGTTTTCGCATCCAAGCCTGCGTTCTATCCGGCATACGGTGGTGGCCGCCGCG
CGGTGCAGGCGGCAAGCATGAAGAAGCGCGTCCTCCACGGCATCCGGCAGCGGCCGTGGGGCAAGTGGGC
GGCGGAGATCCGCGACCCGCACAAGGGCACCCGCGTCTGGCTCGGCACGTTCGACACGGCCGATGACGCC
GCCCGGGCCTACGACGTCGCCGCCCGACGCCTCCGCGGCAGCAAGGCCAAGGTCAACTTCCCCGACGCGG
CCAGGGCCGGGGCGCGCCCGCGCCGCGCCAGCCGTAGAACCGCGCAGAAACCGCCATGCCCCCCTGCGGG
GACGACGGCGTACTCTGCCACCGCAGCATCACGCGCACAGCCGGAGCAGGACACTATGATGGCCAAGCCC
GAGCGGATGGAGTTTTCGGACGTGGACACGTTCGTTGACCTGACCGCCGCCGTCGCCGCGCTACCGCCTG
TCACGGCGAGCTCCTTCGCCGACAAGATGCCGAGGGTCGACGAGGACTCGTCGGAGGGGAGCGGCGGCGG
CGCCATGCTGGGGTTCGCCGACGAGCTTGGGTTCGAACCGTTCATGATGTTCCAGTACGAATCCATGGAC
AGCCTCTTCGCCGGACACGCCGTCATCCAGGATGCCCGCGGTGTGGACGGCGGCATGGACGGCGTTAGCC
TCTGGAGCTTCGACGAGTTCCCCATGGACAGCGCCATTTTC**TGA**CGTTTTCCGTGCGATTTTTTTACGCT
CGATCGTTTGTAAGAATCTCCATCTGGCATCTTGTTCCTTGTAAATATTCGTGCACAGAACCTTGCTCAG
ACCAGATTCTGATTCATGGCCAGGAACCAAGGAAGGGTGCAGCTGATGCATGGGTACTGCATGATTGCTT
CCTTGATGAATGCTGATGGAAATGTATTCTACTGTTTGAGTTTATTGTTCGTCAGTCAGACCCTGCTGTA
TTCTACTGTACTG

>**TaERF4-1** clone 10,1,3 1327bp CDS complete, mRNA 129. 1139 **B-1**

TCCAACTCAACACGCCAAGGCAAAGCACCAGTAGTCACACAAGAGAAAAGCAAGAGCAGCAAGCAGCACA
GAGCTAGAGAAAAGAGGAGCGTCTCATCGGGAGGAGGAGAAAGAAAAAGCGAGCGAGG**ATG**TGTGGCGGA
GCGATCCTCGCCGAGCTCATCCCGGGCGCGCCGGCGCGGCGCGCCACGTCCGGCCATGGCCACGTCTGGC
CGGGCAAGGGCGCCAAGCAGACCAAGGCCGCCGCGGCCGACGACTTCGAGGCCGCGTTCCGGGAGTTCAA
CGAGGACTCTGATGTGGAGGACGACGTCGTGATGGTGATGGAGCGGCGGGAGGAGGTGGCCGAGAGCAAG
CCCTTCGTGTTCGCCGCTTCGCCCAAGAAGCAGCAGCAGGAGGAGGAGGAGGCGGCGCCCGTCCGCCGCA
GGAAGCCGGCGCAGTACCGGGGCGTGCGGCGCCGGCCGTGGGGCAAGTGGGCCGCCGAGATCCGCGACCC
CGTCAAGGGCGTCCGCGTCTGGCTCGGCACCTTCCCCTCCGCCGAGGCCGCCGCGCTCGCCTACGACGAA
GCCGCGCGCGCCATCCGCGGGCCCAGGGCCAAGCTCAACTTCTCCTCCTCCTCTGCCGTCGCCGCCGCGG
CCCCGGGAGCGCGCAAGCGCGGCCGCGCCGCGCCCCCCGCTGCCAAGCCAGCGCCGGTCATCACCCTCGT
CGACGATGAGGAGGAGCACGCGTCGTCCTTCGTCAAGCACGAGGCCGAGGCGAGCGAGGGCTCCGAGTCC
AGCGGCGCCCTCCCCGACTTCTCGTGGCAGGGCATATCGGCGTTCGACGAGGCCCCGGCGTACCCCGCCC
CGGAGCCGGAGACCGAGCAGCTGACAAAGCGGGCGAGGACGACGGAGGCGGAAGACACCGACGAGGGCAT
GTCGGCCCACCCGGCATCAGACTCCGACTCCGACGCGCTCTTCGACGCCCTCCTCTTCGCCGACCAGTTC
GCCTTCTTCAACGGCGGCGCCTACGAGTCCCTCGACAGCCTCTTCAGCGCCGACGCCGTGCAGAGCAGCG
CCACCGCCACCGCCGTGAACGAGGCGGCCCTGGGGCTCTGGACCTTCGACGACGACTGCCTCGTCGACGA
GTGCAGCCTGTCGTTC**TAG**ATCTTGTTCCGTGTGCGTCCGTTCATGTCATGAGATGCGGTGGTGTAAATA
GAGTAGCTTGCCTTGTGATGAAATCCGTTCGTCCGTCGTCGCTTCTTGGATCAGCAAGCGATGATGGATT
CTGTGTAAGAGTTTGGCATCGTCGTGAAAGATTTTGCCTACGCCTTCGATGCAAAAAAAAAAAAAAA

>**TaERF4-2** clone WEF029_J03 1019bp CDS complete, mRNA 211.. 837 **B-1**

GACCATCGAGAGACAGGAGGCGCGGTACTCTCCTAGATAATCCTCGACGCTAACCACGCTTTCCCTTCCC
CTCTCCGAGCCCAAAACCGAGCAAAATCGATTTCAAGCGGACCAAAATCCGGCACCGCATGGGTTTCCCC
AGGCACTGATCCGGCGGGTGGACGCTAATGCGGCGCCGCTGAAGGGATCTGCTGCGCCCGCGGGGTTCAC
ATGGCGGCACAGAACGGCAGCGGCGGCGCGGCGGTGGTCCTCAACGTGTACGACCTGACGCCCATAAACG
ACTACCTCTACTGGTTCGGCCTCGGCATCTTCCACTCCGGGATCGAAGTTCATGGTCTGGAATATGCTTT

TGGAGCCCATGATCTCTCAACCAGTGGGGTTTTTGAGGTTGAACCAAAACGTTGCCCTGGCTATGTCTAT
AGAAGGTCTGTATGGATGGGTACAACTGAGATGTCTAGGGCAGAGTTCCGCTCGTTCGTTCAAACTCTTG
CAGGAAAGTACGATGGGAATACATATCATTTGATTTCAAAGAACTGCAACCATTTTACAGATGATGTCTG
TAAGAACATAACCAAAAAACCAGCCCCTGGATGGGTAAATCGGCTAGCAAGAGTGGGTTACTTTTTCAAT
CGCCTCCTACCAAAAAGCATTCAAGTCTCTACTGTTACTCACGCAGCAACCCATCCTGCATTTTCTGATG
ATGACGTGGATTCATCAGGTTCATCCCTTAATGGGGACAGTGATATGGATGACCTGGACCAGCACCTGCT
ACCAGCCGCAGACACCGACCTTCAATCCATAGTCGTGTCACCAAAGCCATCGAAAGATCTCGTT**TGA**ATT
ATCCTACCATTTGGGTTCCCTGCTGCAGTTTAACGGTCGTACTTCGCCAGCTTGTGTGATATCCAAGTAT
AGTTGAATGTTGATTGTATTGATTCTTTTCCAACGGGTTTTGATGTGTACGCATATAGGAATTGTCGGGT
AAAGAAGTGATTTTGCAAAAAAAAAAAAAAAAAAAAAAA

>**TaEREBP1-1** clone C48R1L1-7 1588bp CDS complete, mRNA 105.. 1301 **B-2**

CTCACACCTGCTCCCCCCCCCTTCATTAAAAACTGCCCATTTATTTACCAATTCCGGCGAACCCCTCCCAC
CGGCGACGACTAGGACGAGGCACCGCCGGCAATC**ATG**TGCGGCGGCGCGATCCTCAAGGGCCTCAAGGTC
CCCGCGCCGGCGCGGAAGGTGACGGCGGCTGTGCTGTGGCCCGAGAAGAACAAGCCCAAGCGGGCCGACG
GCGGCGCCCGGCACCTCGCGGGGCTCGGCCGGCGTGGGGGGCTCGGGCTGGACGACGGCGAGGCGGACTT
CGAGGCCGACTTCGAGGAGTTCGAGGCCGACTCCGGGGACTCCGACCAGGAGCTCGGGCGCGGCGGGGTG
GCTGAGAAGGACGGCGACGACGAGGTCGTCGAGACCAAGCCCTTCGCCGCCGTCAAGAGGTCCCTCTCCC
AAGATGACTTAAGCACCATGACCACTGCTGGTTTTGATGGTCCTGCACAAAGGTCAGCAAAAAGGAAGAG
AAAGAACGAATTCAGGGGTATCCGCCAGCGCCCCTGGGGTAAGTGGGCTGCTGAAATCAGAGATCCTAGC
AAGGGTGTCCGTGTTTGGCTCGGTACCTTCAACAGTGCTGAAGAAGCTGCAAGAGCTTATGATGTTGAAG
CACGAAGGATCCGTGGCAAGAAGGCCAAGGTTAACTTTCCAGAGGAACCAACAGTTCCTCAGAAGCGCCG
TGCTTGCCCTGCTGCTCCTAAAGTTCCCAAGTCAAGCGCAGCACAGGAACCTACCGTCATACCAGCAGTC
AACAACCTTGCCAACCCAAATGCTTTCGTCTACCCGTCTGCTGACTTTGCATCAAAGCAGCCACTTGTTC
AGCCTGATAACGTGCCATTTGTTCCTGCAATTAACTCTGCTGCACCTGTTGAAGCTCCTGTTATGAATAT
GTACTCTGATCAGGGAAGCAACTCCTTTGGCTGCTCTGACTTGGGCTGGGAGTATGACACCAAGACTCCA
GATATATCATCCATTGCTCCCATTTCTACCATTGCTGAAGGAGCAGAAACTGCACTTCTCCAGAGTAACA
CCTACAACCCAGCGGTGATTGCTGAAGGAGCTGAATCTGCGCCTGTCCAGACCAACACCTACAACTCAGT
GGTGCCTCCTGTCATGGAGAACAATGCTGTTGATTTTGAACCTTGGATGAGGTTTCTTTTGGATGATGGC
GTGGATGAGCCGATTGATAGCCTTCTGAATTTTGATGTGCCTCAGGATGTCGTTGGCAACATGGACCTTT
GGAGCTTCGATGACATGCCCAT*CTG*TGGCGAAATTTTC**TGA**GGGATTCAAACCCCTGTACATCAGGACAA
AGGGAATAAGACTATGGGAGATTGGGAAGCACCCTGTTGTCACCTTTGGTTAGAATATGCCTATGTGCAA
GCTTAGATGCAAAAAAGCCGCATCCTGAGTCTCTTTTGTAGTCGACCCTTTCCCTAGTTGACTGTCGATG
AACCGTTGTCATTTATGAGTCCGATGAACCGTTGTCTTTATCTTTTGTCAATTTTTATGTCGTTGCTACC
ATTTCTGAATGTGAACAGCATGGATCTGTGTCCAGTTTGCTTTATCTG

>**TaEREBP1-2** clone C49R1L7-26 1469bp CDS complete, mRNA 81.. 1149 **B-2**

ACACATCATCACACACCACACCAACCGACCCGTCTACTTTCTTTCGCCTTGCTCTACCCGCGGTCCCACG
TGCTCCCAGCC**ATG**TGCGGCGGCGCCATCCTCTCCGACATCATCCCGCCGCCGCGCCGGGCCACCGGCGG
CAACGTCTGGCGGGCGGACAAGAAGAGGAGGGCCAGGCCCGACGCCGCCGCGGGGAGGCCCCGCCGCGTG
CCCGAGGAGGAGTTCCAGGAGGAGGAGGGCGACGCGGAGTTCGAGGCCGACTTCGAGGGGTTCGTGGAGG
CGGAGGAGGAGTCCGACGGCGAGACCAAGCCCTTCCCCGTCCGCAGGACCGGCTTCTCCGGAGATGGACT
GAAGGCAACTGCTGCTGGTGATGATGACCGTGCCTCAGGGTCTGCTAAAAGGAAGAGAAAGAGCCAGTTC
AGGGGCATCCGCCGCCGCCCTTGGGGTAAATGGGCTGCTGAAATAAGAGATCCTCGCAAGGGTGTCCGTG
TCTGGCTTGGCACTTACAACTCTGCTGAGGAAGCTGCCAGAGCCTATGATGTTGAAGCCCGCAGAATTCG
TGGCAAGAAGGCAAAGGTCAATTTCCCAGAAGAAGCTCCCATGGATCCTCAGCAACGCTGCGCTACCTCT
GTGAAGGTGCCCGAGTTCAACACCGAACAGAAGCCAGTACTCAACACCATGGGCAACACAGATGTGTATT
CCTGCCCTGCTGTTGACTACACCTTAAATCAGCAATTTGTGCAGCCTCAGAACATGTCGTTTGTGCCTAC
AGTGAATGCAGTTGAGGCTCCTTTCATGAATTTTTCCTCTGACCAGGGGAGCAACTCCTTTAGTTGCTCA
GACTTCAGCTGGGAGAATGATATCAAGACCCCTGACATAACTTCTGTGCTTGCATCCATTCCCACCTCAA
CTGAGGTCAATGAATCTGCATTTCTCCAGAACAATGGCATTAATTCAACGGTACCTCCTGTGATGGGTGA
TGCTAATGTTGATCTTGCCGACTTGGAGCCATACATGAAGTTCCTGATGGACGATGGTTCAGATGAGTCA
ATTGACAGCATTCTAAGCTGTGATGTACCGCAGGACGTTGTCGGCAACATGGGCCTTTGGACCTTTGATG
ACATGCCCTTGTCTGCTGGTTTCTAC**TGA**GGGAATCGAGGTCGCTGGGTGCCTGTATATATAGACAAAGG
AATAAGTATTCTGGACATCAACAAGTGCTTGTGTCTGGTGCCTCTAGAATCGAGCAGTAGCGACGTCAGT
CTATGGTTATGTCTAGCTTAAATGGTCAGGAGACCTAAGTCTTTTGCAATAGACCTCTGTCTTGTGCCCC

CAGACTATATTATATCTATATATGAAACCAGTATGTGATGGGAACTGCTTATTTTGTATTCCTGTTTCTA
CCTTATTGTAATTGCTACAAGTGGCTGTAAACCTTTTAACTTTGAAGCCAGTGTTTGGTGTGCCGTGCT

>**TaEREBP1-3** clone AP2_41 1486bp CDS complete, mRNA 93.. 1168 **B-2**

TCACCAACCGACCCGTCTCGTCTCGCAAGCgTTTCACTCTTTCCCTTTCCTTTCGCCTCGCTcTACCCG
CGGTCCCACGTGCTCCAGCAGCC**ATG**TGCGGCGGCGCCATCCTCTCcGACATCATCCCGCCGCCGCGCC
GGGCCACCGgcGGCAACGTCTGGCGGGCGGACAAGAAGAGGCGGGCCAGGCCCGACGCCGCCGCGGGGA
GGCCCCGCCGTgCGCCCGAGGAGGAGTTCCAGGAGGagGAGGGCGACGCGGAGTTCGAGGccGACTTCG
AGGGGTTCGTGGAGGCGGAGGAGGAGTCCGACGGCGAGGCCAAGCcCTTCCCCGTCCGCAGgAGCGgCT
TCTCCGGAGATGGATTgAAGGCAACTgCTGCTGGTGATGATGACTGTgCCTCAGGGTCTGCTAAAAGGA
AGAGAAAGAaCCAGTTCAGGGGCATCCGCCGCCGCCCTTGGGGTAAATgGGCTGCTGAAATAAGAgATC
CTCGCAAGGGTGTCCGTGTCTGGCTTGGTACTTACAACTCCGCTGAGGAAGCTGCCAGAGCCTATgATG
TTGAAGCCCGCAGAATTCGTGGCAAGAAGGCAAAGGTCAATTTCCCAGAAGAAGCTCCTATGGCTCCTC
AGCAACGCTGCGCTACTGCtGTGAAGGTGCCCGAGTtCAACACCGAACAGAAGCCGGTACTCAACACCA
TGGGCAACGCAGATGTgTATtCCTGCTCTGCTGTTGACTACACCTTAAATCAGCAATTTGTGCAGCCTC
AGAACATGTCGTTTGTGCCTACAGTgAATgCAGTtGAgGcCCCTtTCATGAATtTTTCCTCTgAcCAGG
GtAGCaACTCCTTTAGTTGCtCAGACTTCAgCtGGGAgAATgAtATCAAGACCCCTGAcATAACTTCTG
TGCTTGCATCCATTCCCACCTCAACAGAGGTCAATGATTCTGCATTTCTCCAGAACAATGGCATCAATT
CAACGGTaCCTCCTGTGATGGGTGATGCTAATGTTGATCTTGCCGACTTGGAGCCATACATGAAGTTCC
TGATGGACGATGGTTCAGATGAGTCAATTGACAGCATTCTAAGCTGTGATGTACCCCAGGATGTGGTCG
GCAACATGGGCCTTTGGACCTTTGATGACATGCCCTTGTCTGCTGGTTTCTAC**TGA**GGGAATCGAGGTC
GCTGGGTGCCTGTATATATAGACAAAGGAATAAGTATTCAGGACATCAACAAGTGCTTGTGTCTGGTGC
CTCTAGAATTGAGCAGTAGCGATGTCAGTCTATGGTTATGTCTAGCTTAAATGGTCAGGTGACTGAGGT
CTTTTGCAATAGACCTCTGTCTTGTGCCCCCAGACTATATTATATCTATATATGAGACCAGTATGTGAT
GGGGAACTGCTTATTTTGTATTCATGTTTCTACCTTATTGTAATTGCTACAAGAGGCTGTAAACCTTTT
AAATTTGAAGCCAGTGTTTGTTATGTTGTGCTTAAAA

>**TaEREBP1-4** clone C55R1L1-20 1578bp CDS complete, mRNA 139.. 1284 **B-2**

ACCATCTCACCACCCTCCTCCCTCCCTCCCTCCCCCACTCCGCCCGCAACTCCTTCACTGTGCCGCCCGT
GCTCTCCCATCCGCTCCGCTCCCCGCGCCGATCCAAAGCCCAGACCCTCGCCTTGATCCGGCCTCGCG**AT**
GTGCGGCGGAGCCATCCTCGCGGGCTTCATCCCGCCGTCGGCGGCCGCCGCGGCGGCCAAGGCGGCGGCA
GCCAAGAAGAAGCAGCAGCAGCGCAGCGTGACGGCCGACTCGCTGTGGCCGGGCCTGCGGAAAAAGGCGG
CCGAGGAGGAGGACTTCGAGGCCGACTTCCGCGACTTCGAGCGGGACTCCAGCGACGACGACGCCGTGGT
CGAGGAGGTTCCACCGCCGCCGGCCTCGGCGGGTTTCGCCTTCGCCGCCGCGGCCGAGGTCGCGCCCCCG
GCCCCTGCCCGCCTAGATGCTGTTCAACATGATGGACCTGCTGCCAAACAAGTAAAGCGCGTTCGGAAGA
ATCAGTACAGAGGGATCCGCCAGCGTCCCTGGGGGAAATGGGCAGCTGAAATCCGTGACCCTAGCAAGGG
TGTCCGGGTTTGGCTCGGGACATACGACACTGCTGAGGAGGCAGCAAGGGCATATGATGCTGAAGCCCGC
AAGATTCGTGGCAAGAAGGCCAAGGTCAATTTTCCTGAGGATGCTCCAACTGTTCAGAAGTCTACTCTGA
AGCCAACTGCCGCTAAATCAGCAAAGCTGGCTCCACCTCCGAAGGCCTGCGAGGATGAGCCTTTCAATCA
TCTGAGCAGAGGAGACAATGATTTGTTCGCGATGTTCGCCTTCAATGACAAGAAAGTTCCTGCGAAGCCA
GCTGAAAGTGTGGATTCCCTTCTTCCGGTGAAACCTCTTGTGCCCACTGAGACATTCGGGATGAACATGC
TCTCTGACCAGAGTAGCAACTCATTTGGCTCTACTGACTTTGGGTGGGACGATGAGGTCATGACCCCGGA
CTACACGTCCGTCTTCGTCCCGAATGCTGCTGCCATGCCGGCATACGGCGAGCCCGCTTACCTGCAAGGT
GGAGCTCCAAAGAGAATGAGGAACAACTTTGGCGTGGCCGTGCTGCCTCAGGGAAATGTTGCACAAGACA
TCCCTGCTTTTGACCATGAGATGAAGTACTCGTTGCCTTATGTTGAGAGCAGCTCGGACGGATCAATGGA
CAGCCTTCTGCTGAATGGTGCGATGCAGGACGGGGCAAGCAGTGGGGATCTCTGGAGCCTCGATGAGCTC
TTCATGGCGGCTGGTGGTTAC**TGA**GGGTTCTTGTCTGTGTGGTCTGCGGATAGCACAAATGTCCCTTGCA
TGTGGCCAAGACGAAGAACTGGTGGTGCATGTGGCCAAGATGAAGGATTGATTGCATCTGCTATGTTTCG
TAGCCGGATCAAAACCTAGTTATGCTGAAGACTGTATGCTGCTAGCAGTGGAACCGTATGTCATGTTTAT
AAGTATTTTGTTGTTGTACATCGCCTCTATGATTGGGTGCATGTTGGAGACTGGAGTTTAATAAATAAAT
ACCTTGGTCATATGCCTGATGTAATGTGTGTGTTGAGT

>**TaEREBP1-5** clone c55f117-44 1523bp CDS complete, mRNA 90.. 1232 **B-2**

CTTCTTCACTGTGCCGCCCGTGCTCTCCCATCCGCTCCGCTCCCCGCGCCGATCCAAAGCCCAGACCCTC
GCCTTGATCCGGCCTCGCG**ATG**TGCGGCGGAGCCATCCTCGCGGGCTTCATCCCGCCGTCGGCGGCCGCC

GCGGCGGCCAAGGCGGCGGCAGCCAAGAAGAAGCAGCAGCAGCGCAGCGTGACGGCCGACTCGCTGTGGC
CGGGCCTGCGGAAAAAGGCGGCCGAGGAGGAGGACTTCGAGGCCGACTTCCGCGACTTCGAGCGGGACTC
CAGCGACGACGACGCCGTGGTCGAGGAGGTTCCACCGCCGCCGGCCTCGGCGGGTTTCGCCTTCGCCGCC
GCGGCCGAGGTCGCGCCCCCGGCCCCTGCCCGCCTAGATGCTGTTCAACATGATGGACCTGCCAAACAAG
TAAAGCGCGTTCGGAAGAATCAGTACAGAGGGATCCGCCAGCGTCCCTGGGGGAAATGGGCAGCTGAAAT
CCGTGACCCTAGCAAGGGTGTCCGGGTTTGGCTCGGGACATACGACACTGCTGAGGAGGCAGCAAGGGCA
TATGATGCTGAAGCCCGCAAGATTCGTGGCAAGAAGGCCAAGGTCAATTTTCCTGAGGATGCTCCAACTG
TTCAGAAGTCTACTCTGAAGCCAACTGCCGCTAAATCAGCAAAGCTGGCTCCACCTCCGAAGGCCTGCGA
GGATGAGCCTTTCAATCATCTGAGCAGAGGAGACAATGATTTGTTCGCGATGTTCGCCTTCAATGACAAG
AAAGTTCCTGCGAAGCCAGCTGAAAGTGTGGATTCCCTTCTTCCGGTGAAACCTCTTGTGCCCACTGAGA
CATTCGGGATGAACATGCTCTCTGACCAGAGTAGCAACTCATTTGGCTCTACTGACTTTGGGTGGGACGA
TGAGGTCATGACCCCGGACTACACGTCCGTCTTCGTCCCGAATGCTGCTGCCATGCCGGCATACGGCGAG
CCCGCTTACCTGCAAGGTGGAGCTCCAAAGAGAATGAGGAACAACTTTGGCGTGGCCGTGCTGCCTCAGG
GAAATGTTGCACAAGACATCCCTGCTTTTGACCATGAGATGAAGTACTCGTTGCCTTATGTTGAGAGCAG
CTCGGACGGATCAATGGACAGCCTTCTGCTGAATGGTGCGATGCAGGACGGGGCAAGCAGTGGGGATCTC
TGGAGCCTCGATGAGCTCTTCATGGCGGCTGGTGGTTACTGAGGGTTCTTGTCTGTGTGGTCTGCGGATA
GCACAAATGTCCCTTGCATGTGGTCAAGACGAAGAACTGGTGGTGCATGTGGCCAAGATGAAGGATTGAT
TGCATCTGCTATGTTTCGTAGCCGGATCAAAACCTAGTTATGCTGAAGACTGTATGCTGCTAGCAGTGGA
ACCGTATGTCATGTTTATAAGTATTTTGTTGTTGTACATCGCCTCTATGATTGGGTGCATGTTGGAGACT
GGAGTTTAATAAATAAATACCTTGGTCATATGCCTGATGTAAAAAAAAAAAAA

>**TaEREBP1-6** clone C57R1L1-1 1758bp CDS complete, mRNA 353.. 1560 **B-2**

TTGCGGCCGCGAAATTCGCCCTTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATTAGGCCT
ATTTAGGTGACACTATAGAACAAGTTTGTACAAAAAAGCAGGCTGGTACCGGTCCGGAATTCCCGGGATA
TCGTCGACCCACGCGTCCGCGACTGGAGCACGAGGACACTGACATGGACGAAGGAGTAGAAAACCATCTC
ACCACCACCCTCCTCCCTCCCTCCCTCCCCCACTCCGCCCGCAAAGCTCCCTCGCCTCCTCCTTCACTGT
GCCGCCCGTGCTTCCCCTCCGCTCCCCGCGCCGATCCAAAGCCCAGACCCTCGCCTTGATCCGCATCTCG
CCATGTGCGGCGGAGCCATCCTCGCGGGCTTCATCCCGCCGTCGGCGGCCGCCGCGGCGGCCAAGGCGGC
GGCAGCCAAGAAGCAGCAGCAGCAGCAGCAGCGCAGCGTGACGGCCGACTCGCTCTGGCCGGGCCTGCGG
AAGAAGCCGGCCGAAGAGGAGGACTTCGAGGCCGACTTCCGCGACTTCGAGCGGGACTCCAACGACGACG
ACGACGCGGTCGAGGAGGTCCCCCCACCGCCGGCCACGGCGGGCTTCGCCTTCGCCGCCGCGGCCGAGGT
CGCGCTCCCGGCTCCGACCCGCCTAGATGCTATTCAACATGATGGACCTGCTGCCAAATCAGTGAAGCGC
GTTCGGAAGAATCAGTACAGAGGGATCCGCCAGCGTCCCTGGGGGAAATGGGCAGCTGAAATCCGTGACC
CTAGCAAGGGTGTCCGGGTTTGGCTCGGGACATACGACACTGCTGAGGAGGCAGCCAGGGCATATGACGC
TGAAGCCCGCAAGATCCGTGGCAAGAAGGCCAAGGTCAATTTTCCTGAGGAGGCTCCAACTGTTCAGAAG
TCCACCCTGAAGCCAACTGCTGTGAAATCAGCAAAGCTGGCTCCACCTCCGAAGACCTGCGAGGATGAGC
CCTTCAATCACCTGAGCAGAGGAGACAATGATTTGTTCGCGATGTTTGCCTTCAATGACAAGAAGGTTTC
TGCAAAGCCAGCTGAAAGTGTGGATTCCCTTCTTCCAGTGAAACCTCTTGTGCCCACTGAGACATTCGGG
ATGAACATGCTCTCTGACCAGAGCAGCAATTCATTTGGCTCCACTGACTTTGGGTGGGACGACGAGGCCA
TGACCCCAGACTACACATCCGTCTTCGTTCCGAATGCTGCTGCCATGCCGGCATACGGCGAGCCCGCTTA
CCTGCAAGGTGGAGCTCCAAAGAGAATGAGGAACAACTTTGGTGTAGCCGTGCTGCCTCAGGGAAATGGT
GCACAAGACATCCCTGCTTTTGACCATGAGATGAAGTACTCGTTGCCTTATGTCGAGAGCAGCTCGGACG
GATCGATGGACAGCCTTCTGCTGAATGGTGCGATGCAGGACGGGGCAAGCAGTGGGGATCTCTGGAGCCT
TGATGAGCTCTTCATGGCGGCTGGTGGTTATTGATGGTTCTTGTCAGTGTGGTCTGCGGATAGCACAAAT
GTCCCTTGCATGTGGCCAAGATGAAGAAGTGGTGGTGCATGTGGCCAGGATGAAGGATAGGTTGCATCTG
TTATGCTTGGTAGCGGATCAAACCCTAGCTATGCTAAAGACTGTATGCTGCTAGCAGTGGAACCGTATGT
CATGTTTATAAGTATTTTGTTGTTGTATATCGCCTCTATGATTGGGTGCATGTGGAAGGGCGAATTCGTT
AACCTCAG

>**TaEREBP1-7** clone C94R1L1-1 1222bp CDS complete, mRNA 94-970 **B-2**

GACTCCACACCCTCGGTTCACACACAGTGATCAGCTCAGCTACAGCCTACAGAGCACAGACCCAACCCCG
GGAGACAACAGAGCACGGCGAGGATGCTGCTGAACCCGGCCTCGGAGGCGATGGTGCTGGACAGCATCAG
GCAGCACCTCATGGAGGACACGGCGGCTCCGGCGCCGGAGGCGAGGAGGCAGAGGCCGGCCTACTGCCGC
AGCGCCAGCTTCGGGAGCCTGGTGGCCGACCAGTGGAGCGAGTCGCTCCCGTTCCGCGCGGACGACTCCG
ACGACATGGTCGTCTACGGCGCGCTCCGCGACGCCTTCTCCTGCGGCTGGCTTCCCGACGGCTCCTTCGC
CGCCGTCAAGCCGGAGCCCCTCCTGCCTTCCCCGGACTCCTGCTATGTCGGCTTCCTCGAGCCGGAGACG

CCTCCGGCCACGTCCCCGGGCGGGAAGCGAGGAGGAGGCCGCCGCGGCGTTTCATGGGCCGAGGAGGCGG
CCGCGGCCGTGGCGAGGGGGAAGCACTACCGCGGGGTGAGGCAGCGGCCGTGGGGCAAGTTCGCGGCGGA
GATCCGGGTCCCGGCCAAGAACGGCGCCCGCGTGTGGCTCGGCACCTACGACACCGCCGAGGACGCCGCC
GTGGCCTACGACCGCGCCGCGTACCGCATGCGCGGCTCCCGCGCGCTCCTCAACTTCCCGCTCCGCATCG
GCTCCGAGATCGCCGCCGCGCACGCCGCGGCCGCGGCCGCCGGCGAACAAGCCGCCGTCGCCCGAGCCCG
CGACATCGGACTCGTCATCCTCCTCGTCCTCCTCCGGGTCCCCGAAAAGGAGGAAGAGAGGCGAGGCCGC
GGCCGCGTCCATGGCCATGGCGCTGGTGCCGCCGTCCTCGCAGTTGAGCCGGCCGGCCCACGCGTGGTAC
CCCGCCGTGCCGGCGGAGCAGGTGGCCATGGCGCCGTGCGTGCAGCAGCTGGTCAGC**TAG**GCCCCGCTCT
CCAGCCCCCGGGCGCACGATCGTGAAAGACAATGCGCAAGTGCAGTGGATTCGCCAGCAAGTGCATGGAC
GCGACAGCAACAGGAGCCCGCCTTTCACTGTTGCACGGTGCAAGGTGGTGATGGCAATGGAGTGCGATAC
CTATCGAGCGAGCAGGCAAAGCAACCTTGGCGACGGCTCCATTTGGATAGATATTTGGATATTTGATGGC
AATGGAAGATAAGCTCTCGGAGCAACTGGGCG

>**TaEREBP1-8** clone C95R1L1-8 1381bp CDS complete, mRNA 94-1014 **B-2**

ATTCGACTCGGCGCCCCCGGTTCACACACACTGATCGGCTCAGCTACAGCCTACGGAGTACAGACCCAAC
CCAACCCCGGGAGACAACAGGGC**AtG**GCGAGGATGCTGCTGAACCCGGCCTCGGAGGCGCTGGTGCTCGA
CAGCATCCGGCAGCACCTCATGGAGGACACGGCGGCCGCGGCGGCGCCGGTGGCGGAGGCGAGGCGGCAG
CGGCCGGCCTACTGCCGCAGCGCCAGCTTCGGGAGCCTGGTGGCGGACCAGTGGAGCGAGTCGCTCCCGT
TCCGCGCCGACGACTCCGACGACATGGTCGTCTACGGCGCCCTCCGCGACGCCTTCTCCTGCGGCTGGCT
CCCCGACGGCTCCTTTGCCGCCGTCAAGCCGGAGCCCCTCCTGCCTTCCCCGGACTCTGCCTACGCGTGC
GACGGCCCGTCCTGCTTCGGTTTCCTCGAGCCGGAGACGCCCCCCGCCACGTCCCCGGGCGAGGGCGAGG
AGGAGGGCGCGGCGTTCATGGGCGAGGCGGCCGCCGCGGTGGCGAGGGGCAAGCACTACCGCGGGGTGCG
GCAGCGGCCGTGGGGCAAGTTCGCGGCGGAGATCCGGGACCCGGCCAAGAACGGCGCGCGCGTGTGGCTC
GGCACCTACGACACCGCCGAGGACGCCGCCGTGGCCTACGACCGCGCCGCGTACCGCATGCGCGGCTCCC
GCGCGCTCCTCAACTTCCCGCTCCGCATCGGCTCCGAGATCGCCGCCGCGCACGCCGCGGCCGCGCACGC
CGCGGCCGCGGCCGCCGGGGACAAGCGGCCGTCACCCGAGCCCGCGACCTCGGACTCGTCATCCTCGTCC
TCCTCCTCCGGGTCCCCGAAAAGGAGGAAGAGGGGCGAGGCCGCGGCCGCGTCGATGGCCATGGCGCTGG
TGCCGCCGCCCTCGCAGCTGAGCCGGCCAGCCCACGCGTGGTACCCCGCCGCGCCGGTGGAGCAGGTGGC
CATGGCGCCGCGCGCGCAGCAGCTGGTCAGC**TAG**GCCCCGCCCGCCAGCCCCCGGGCGCACGATCGTGGA
AGACAATGCGCAAGTGCAATGGATTGGCCAGCAAGTGCATGGACGCGACAGCAACAGGAGCCCGCCTTTC
ACTGGTGCACGGTGCAAGGTGGTGATGGCAATGGAGTGCGATGCCTATCGAGCGTGCGAGCGAGCAATCG
GGCAAGCAAAACCTCGGCGACGGTCCCATTGGGATAAATATTTGGATGGGCCGAAAAAGGAAAAAACCTC
GGACCAACTGGGCGCTGCAATTTGGCCTGGAATGAAAAAAAGTCCCTTCCCCCCTGGTTCGGGGAACTCC
CCCTAAGGGATCAAAACGGAACCCGGCTAATATCGGGCCCCCCGGGTTCCT

>**TaEREBP1-9** clone C124R1L1-18 1438bp CDS complete, mRNA 338. 1366 **B-2**

AAAGGGGGAAATTGAATTTAGGGGCCCGGAAATTCCCCCCTGAGGGGGGAAACCATTTTCCCCCCGGGAA
CCCCCCTTTTGGCCCCCTTTGGGCTATTTAGGTGCCACTATAGAACAAGTTTGTACAAAAAAAGCAGGCT
GGTACCGGTCCGGAATTCCGGGGATATCGTCGACCCACGCGTCCGCGACTGGAGCACGAGGACACTGACA
TGGACTGAAGGAGTAGAAAACGCAAAGCAGAGTCCATTCCACTAGTCGAGAGAAAATCAAGAGCAGCGAC
CAGAGCTAGGGGGGAAAAGCAGAGGCAAAGCATCGAAAGAAGAGAAAGAGAGCGAGG**ATG**TGTGGCGGAG
CGATCCTCGCCGAGCTCATCCCGGGCGCGCCGGCTCGCCGCGCCGCGTCCGGCCACGGCCACGTCTGGCC
CGGCAAGGGCGCCAAGCAGACAAAGGCCGCGGCGGCCGACGACTTCGAGGCCGCGTTCCGGGAGTTCAAC
GAGGACTCTGATGAGGAGGACGTGGTGATGGTGGTGGAGCGGCAGGAGGAGGTGGCCGAGAGCAAGCCCT
TCGTGTTCGCCGCTTCGCCCAAGAAGAAGCAGCAGCAGCAGCAGGAGGAGGAGGAGGAGCAGGCGGCGCC
CGCCCGCCGCAGGAAGCCGGCGCAGTACCGGGGCGTGCGGCGCCGGCCGTGGGGCAAGTGGGCCGCCGAG
ATCCGCGACCCTGTCAAGGGCGTCCGCGTCTGGCTCGGCACCTTCCCCTCCGCCGACGCCGCCGCGCTCG
CCTACGACGAGGCCGCGCGCGCCATCCGCGGGCCCAGGGCCAAGCTCAACTTCCCTTCCTCCACCGCCGT
CGCCGCCGCGGCCCCGGGAGCACGCAAGCGAGGCCGCGCCGGCGCCGACGCGGCCCCCGCTGCCAAGCCA
GCGCAGGTCATCACCCTCGTCGACGAAGAGGAGGAGCACGCGTCGTCCTTGTTCGTCAAGCACGAGGCCG
ACGCGAGCGAGGGCTCCGAGTCCAGCGGCGCCCTCCCCGACTTCTCGTGGCAGGGCGTGTCGGCGTTCGA
CGAGGCCCCGGCGTACCCCGCCCCGGAGCCCGAGACCGAGCAGCTCGCGAAGCGGGCGAGGACGACGGAG
GCCGAGGACGCCGACGAGGGCATGTCGGCCCACCAGGCGTCCGACTGCGACTCCGACGCGCTGTTCGACG
CGCTCCTCTTCGCCGACCAGTTCGCCTACTTCAACGGCGGCGCCTACGAGTCCCTCGACAGCCTCTTCAG
CGCCGACGCCGTGCAGAGCGGCGCCACCGCCGCGGACGAGGCGGCCCTGGGGCTCTGGACCTTCGACGAC
GACTGCCTCGTCGACGAGTGCAGCCTGTCGTTC**TAG**ATCTTCTTGTTCCGCGTGCGTCCGTTCATGTCAT

GCGATGCGGTGGTGAAAAAGGGCGAATTGGTTAACCTC

>**TaEREBP1-10** clone BQ237319R1L8-31 1798bp CDS complete, mRNA 492. 1562 **B-2**

CAGGGTTCGTTTTCCGGAAAGGAAACCGGGGACGGGAAACCCAAATAAATTAAATTGGGTTTAATCCCTT
GGGTTCGCCCGGCCAGGGCTTTAAAATTTAAGGTTCCGGATTCTTATGGGAATTGGAAACGGAAAGCATT
TCATTTCGGGCAACGGAATAGGGCCGGGCTAGGTCAGGCTAAGATTAAACTTACCCTACAGGAAAGGGTC
TAGTCCGGTTAAATTAAATTGACCTTGCCCGGAAAGCATTTTCACTTCGGGACACGGTTAGCCCATTACG
CTTGTCTAGTTGACGTTACAGTACAGGTTAGTCAAAAAAGCAGGGCGGTGCGGTTCCGGTCTTCGCTGGC
TATGATATAGTCGAGCCGTCGGATCATCACACACACACACCAACACCACACCAACCGACCCGTCTCGT
CTCGCAAGCGTTTCACTCTTTCCCTTTCCTTTCGCCTCGCTCTACCCGCGGTCCCACGTGCTCCAGCAGC
C**ATG**TGCGGCGGCGCCATCCTCTCCGACATCATCCCGCCGCCGCGCCGGGCCACCGGCGGCAACGTCTGG
CGGGCGGACAAGAAGAGGCGGGCCAGGCCCGACGCCGCCGCGGGGAGGCCCCGCCGTGCGCCCGAGGAGG
AGTTCCAGGAGGAGGAGGGCGACGCGGAGTTCGAGGCCGACTTCGAGGGGTTCGTGGAGGCGGAGGAGGA
GTCCGACGGCGAGGCCAAGCCCTTCCCCGTCCGCAGGAGCGGCTTCTCCGGAGATGGATTGAAGGCAACT
GCTGCTGGTGATGATGACTGTGCCTCAGGGTCTGCTAAAAGGAAGAGAAAGAACCAGTTCAGGGGCATCC
GCCGCCGCCCTTGGGGTAAATGGGCTGCTGAAATAAGAGATCCTCGCAAGGGTGTCCGTGTCTGGCTTGG
TACTTACAACTCCGCTGAGGAAGCTGCCAGAGCCTATGATGTTGAAGCCCGCAGAATTCGTGGCAAGAAG
GCAAAGGTCAATTTCCCAGAAGAAGCTCCTATGGGTCCTCAGCAACGTTGCGCTACTGCTGTGAAAGGTG
GCCCGAGTTCAACACCGAACAGAAAGCCGGTACTCAACACCATGGGCAACGCAGATGTGTATTCCTGCTC
TGCTGTTGACTACACCTTTAATCAGCAATTTGTGCAGCCTCAGAACATGTCGTTTGTGCCTACAGTGAAT
GCAGTTGAGGCCCCTTTCATGAATTTTTCCTCTGACCAGGGTAGCAACTCCTTTAGTTGCTCAGACTTCA
GCTGGGAGAATGATATCAAGACCCCTGACATAACTTCTGTGCTTGCATCCATTCCCACCTCAACAGAGGT
CAATGAATCTGCATTTCTCCAGAACAATGGCATCAATTCAACGGTACCTCCTGTGATGGGTGATGCTAAT
GTTGATCTTGCCGACTTGGAGCCATACATGAAGTTCCTGATGGACGATGGTTCAGATGAGTCAATTGACA
GCATTCTAAGCTGTGATGTACCCCAGGATGTGGTCGGCAACATGGGCCTTTGGACCTTTGATGACATGCC
CTTGTCTGCTGGTTTCTAC**TGA**GGGAATCGAGGTCGCTGGGTGCCTGTATATATAGACAAAGGAATAAGT
ATTCAGGACATCAACAAGTGCTTGTGTCTGGTGCCTCTAGAATTGAGCAGTAGCGATGTCAGTCTATGGT
TATGTCTAGCTTAAATGGTCAGGTGACTGAGGTCTTTTGCAATAGACCTCTGTCTTGTGCCCCCAGACTA
TATTATATCTATATATGAGACCAGTATGTGATGGGGAACGGCATATTT

>**TaERF1-1** clone C55R1L1-39 1335bp CDS complete, mRNA 115.. 1013 **B-3**

TCCTCCTCCCCCATTCCGCCCGCAACTCCTTCACTGTGCCGCCCGTGCTCTCCCATCCGCTCCGCTCCCC
GCGCCGATCCAAAGCCCAGACCCTCGCCTTGATCCGGCCTCGCG**ATG**TGCGGCGGAGCCATCCTCGCGGG
CTTCATCCCGCCGTCCGCGGCCGAGGTCGCGCCCCCGGCCCCTGCCCGCCTAGATGCTGTTCAACATGAT
GGACCTGCTGCCAAACAAGTAAAGCGCGTTCGGAAGAATCAGTACAGAGGGATCCGCCAGCGTCCCTGGG
GGAAATGGGCAGCTGAAATCCGTGACCCTAGCAAGGGTGTCCGGGTTTGGCTCGGGACATACGACACTGC
TGAGGAGGCAGCAAGGGCATATGATGCTGAAGCCCGCAAGATTCGTGGCAAGAAGGCCAAGGTCAATTTT
CCTGAGGATGCTCCAACTGTTCAGAAGTCTACTCTGAAGCCAACTGCCGCTAAATCAGCAAAGCTGGCTC
CACCTCCGAAGGCCTGCGAGGATGAGCCTTTCAATCATCTGAGCAGAGGAGACAATGATTTGTTCGCGAT
GTTCGCCTTCAATGACAAGAAAGTTCCTGCGAAGCCAGCTGAAAGTGTGGATTCCCTTCTTCCGGTGAAA
CCTCTTGTGCCCACTGAGACATTCGGGATGAACATGCTCTCTGACCAGAGTAGCAACTCATTTGGCTCTA
CTGACTTTGGGTGGGACGATGAGGTCATGACCCCGGACTACACGTCCGTCTTCGTCCCGAATGCTGCTGC
CATGCCGGCATACGGCGAGCCCGCTTACCTGCAAGGTGGAGCTCCAAAGAGAATGAGGAACAACTTTGGC
GTGGCCGTGCTGCCTCAGGGAAATGTTGCACAAGACATCCCTGCTTTTGACCATGAGATGAAGTACTCGT
TGCCTTATGTTGAGAGCAGCTCGGACGGATCAATGGACAGCCTTCTGCTGAATGGTGCGATGCAGGACGG
GGCAAGCAGTGGGGATCTCTGGAGCCTCGA**TGA**GCTCTTCATGGCGGCTGGTGGTTACTGAGGGTTCTTG
TCTGTGTGGTCTGCGGATAGCACAAATGTCCCTTGCATGTGGCCAAGACGAAGAACTGGTGGTGCATGTG
GCCAAGATGAAGGATTGATTGCATCTGCTATGTTTCGTAGCCGGATCAAAACCTAGTTATGCTGAAGACT
GTATGCTGCTAGCAGTGGAACCGTATGTCATGTTTATAAGTATTTTGTTGTTGTACATCGCCTCTATGAT
TGGGTGCATGTTGGAGACTGGAGTTTAATAAATAAATACCTTGGTCATATGCCTGATGTAATGTGTGTGT
TGAGT

>**TaERF1-2** clone wef45f12 1029bp CDS complete, mRNA 4.. 652 **B-3**

TCA**ATG**TCTGATCCAAGCAGCACAGCTTCCTCTTATTCCACGTCACCCCGGCTCACCACCGGCGTCGTCA
ACTTCTTGGCGCGCCGCGCCATGACCACGCAGAACCACCACAGAGCAGCAGCATCCCTCCACTCGCCAGG

CTCCTCCACCGGCTCCGCCGACACTGCGCCATGGCACCACCGCGCCCCCGGCCACGCCGCCGCCTCTGCTC
CCGTTCCACGCCAGCGACGCCGACGAGATGCTGCTGCTCGACATGCTCTCCCAGCACCAGGAGGACATGC
ACACCGACACCGCGACAGCGCCCGTTCCCACCGCGACGGCGGCAACGGCCGTGAAGCGAGAGGTCAGCGA
AGAGGAGGAGGCCAAGGTGGCCGCGGGCGGCAGTAGGCGCGCGTTCCGCGGGGTGCGGAAGCGGCCGTGG
GGCAAGTTCGCGGCGGAGATCCGGGACTCGACGCGGGACGGCGTCCGGGTGTGGTTGGGCACGTTCGATA
GCCCGGAGGCGGCGGCGCTCGCGTACGACCAGGCCGCCTTCGCCATACGGGGCGGCGCCGCCGTGCTCAA
TTTCCCCGCCGACCAGGTCCGGCGCTCGCTCGAGGGCGCAGAGGACGACGTGTGTGGCCGCGCCGACGGG
TTGTCACCCGTGCTCGCGC**TGA**AGCGGCGCCACTCCATGCATAGGAGGGCGGCGACAAAACAAAAGGTCA
GGACCGTACGGACGGGCCGGCCGGAGGGCGTGATGGAGCTTGAGGACCTCGGCGCAGAATACCTCGAGGA
GCTGCTCGGCCTCTCCGACGACATGATGACGTCTGCCTCCAGTTCATGGTGCTCGAGTCATCACTCCATC
TGACGATGGATTGATCGCTTTTTTTTTGCTTTCTTGTCTGTTGTTGTTGCTGGTCTGGTCTCGATAGTCC
AGTGTTCAGTTGGACTAGATAGCATTGAGAAATGATAGTGTAAATCGTTGGTTGTGCACGTTTTTAAGCG
TCAATGTTTCGCTTTGAATTCAGTCAAAACAAAAAAAAAAAAAAAAAAA

>**TaERF1-3** clone C56F1L1-8 1455bp CDS complete, mRNA 88..1221 or RAP2.12 **B-3**

GCAGAGCTCCCGTCCTCCCTCCACTGTGCCGCCCGTGCTCCCGGCGCCGATCCGGCCCTCCCCCCCTCCC
TTGATCCACACCTCGCC**ATG**TGCGGCGGAGCCATCCTCGCGGGCTTCATCCCGCCGTCGGCGGCCGCCAA
GGCGGCGGCGGCCAAGAAGAAGCAGCAGCAGCGCAGCGTGACGGCCGACTCGCTCTGGCCGGGCCTGCAG
AAGAAGGCGGCGGACGAGGAGGACTTCGAGGCCGACTTCCGCGAATTCGAGCGGGACTCCAGCGACGACG
ACGCCGCGGTCGAGGAGGTCCCCCCGCCGCCGGCCGCGGCGGGGTTCGCCTTCGCCGCCGCCGCCGAGGT
CGCGCCGCCGGCCCCTGCCCGCCTAGATGCTGTTCAACGTGATGGACCTGCTGCCAAACAAGTAAAGCGC
GTTCGGAAGAATCAGTACAGAGGGATCCGCCAGCGTCCCTGGGGAAAATGGGCAGCTGAAATCCGTGACC
CTAGCAAGGGTGTCCGGGTTTGGCTCGGGACATACGACACTGCTGAGGAGGCAGCCAGGGCATATGATGC
TGAAGCCCGCAAGATCCGTGGCAAGAAGGCCAAGGTCAATTTTCCTGAGGATGCTCCAACTGTTCAGAAG
TcCCACCCTGAAGCCAACTGCTGCtTAAATCAGCAAAGCTGGCTCCACCTCCGAAGGCCTGTGAGGATGG
CCTTCAATCATGTGAGCAGAGGAGACAATGATCTGTTTGCGATGTTCGCCTTCAATGACAAGAAGGTTCC
TGCAAAGCCAGCTGAAAGTGTGGATTCCCTTCTTCCGGTGAAACTGAAACCTCCTACTGAGACATTCGGG
ATGAACATGCTCTCTGATCAGAGCAGCAACTCGTTTGGCTCTACTGACTTTGGGTGGGACGACGAGGCCA
TGACCCCAGACTACACATCCGTCTTCGTTCCGAATGCTGCCATGCCGGCGTATGTCGAGCCCGCTTACCT
GCAAGGTGGAGCTCCAAAGAGACTGAGGAACAACTTTGGCGTGGCCGTGCTGCCTCAGGGAAATGGTGCA
CAAGACATCCCTGCTTTTGACCATGAGATGAAGTACTCGTTGCCTTATGTCGAGAGCAGCTCGGACGGAT
CAATGGACAGCCTTCTGCTGAATGGTGCGATGCAGAACGGGGCAAGCAGCGGGGATCTCTGGAGCCTCGA
CGAGCTCTTCATGGCGGCCGGTGGTTAC**TGA**GGGTTCTTTGTCAGTGTGGTCTGCGAATAGCACAAATGT
CCCTTGCATGTGGCCAAGATGAAGAACTGGTGGTGCATGTGGCCGAGATGAAAGGACTGGTTGCATCTGC
TATGTTTCGTAGCCAGATCAGACCCTAGCTATGGTAAAAGACTTCATGCTGCCAGCAGCGGAACCGTATG
TCATGTTTATAAGTATTTTCTTGTTGTACATCGTCTCTATGATTGGGTGCATGTG

>**TaERF1-4** clone C75R1L1-57 1803bp CDS complete, mRNA 475.. 1515 **B-3**

CTTTTTTTTTTTTTGCTCCTCTCGCCGCTTTGGGTCGGGGTTTTAGCGAGCTCAGGGAGGCCGGCCGGAG
GCGACGACCGCCGTCGGGGAAGAAGAAGAAGACGCCCCACCACTCGTATATGTCTTCTTGCCCCTTCGCC
GATCTGGTGGTAGTCGCGCTGCTCCTCGCCGCCGCCGCCGTCGTCGCCGTCGTCCTCCTCCACGCCCTCG
CCTGCGCCCCCCGATCTATGCACCAGGGGCTCCACGGCTGGAACAAGTCCTCGTCGATGCTCAGGGACGG
GTTCGGGGTGAAGTATTCCGAATTCCTCCACATAAGGCCGTGTAGTTTCTTGCGAGGAGACTGACGGATT
CCCGGTGATTCCGTCCGTCCAATTTTTGACACCCAAGAGAGAGAAAAAGTTGTCCCCTTTTCCTAGTAGA
GTTTTTCCCCTTCAGTTTTGCTAGTCAGTCAGTGTTGGAAGAAGAAGACGATCC**ATG**GCCCCGCTGAGCG
AGCACCAGATGCGCCTCACGCGCGCCCTGGCCAAGAGGCCCAAGACCAAGATGCTCACCGGGTTCGGCGT
CAAACCCACCGCTGCTTTCTCCAAGCCGGCGCAGCTGCAGGCGCAGGCGCTGCTGCCCGGTGCGCAGCCG
CCGCGGCGGCGCGTGCGCGTCCTGTACGAGGACCCCGACGCCACCGACTCGGACACGGACGACGAGGAGG
CTGTCGCCGCCCCTGCTTCGTCCAAGCGCTGCTTCGAGCTGTTCCTTGGCAAGGCCCCGCCGGCCAAGGT
GTTTGCCAAGCCGGTCACTCCGACCGCCGCGGCTTCTGCCGTCGCTGCCTGCACCACCAGCAGCGCCGAG
GGCCACCGTGGGGTGCGCCTCCGCAAGTGGGGCAAATGGGCGGCCGAGATCCGCAACCCGTTCTCCGGCA
AGAGGGAGTGGCTTGGCACCTTCGACACTGCCGACCTGGCCTCCGCCGCCTACCAGGCCGCCTCCCGGAG
CTTTATCGAGGAGAAGCGTCGCCGCCGTGGCCAGTCTGTGGCTGCGGCCTCACCTGCTCGGTCGGCTGCG
TCAACAACACCGACGTCGTCTTCTACTACACCAACGGCATCTTCGTCGTCCTCCACCTCGGCTGCTCCAT
TCGCACACCCGTCGCCGTCCTCTGTCCTGGAGGCCACCAAGCCAGCAGCCGAGTCCCTGTCGCCTGAGCC
GACCCCAGTTCCCGTCATGTTGTCCACCACGGCTGAGACTGCGCAGCTTCCTGACGACCCAGAGTTCTAC

CAGGACCTCCTGCGTGGTCTCCAGCTGCCGGACATCGACCCGATGGACTTCCGCGCCGGGCTTGATGCCT
TGGACGTCTCCGAGGCGTCGGCTTACTTGGACGGCGAACAGGATCTGCTGCTCGGTGACCTGGCCGACGA
GGACCTCGAGCTGGACATGGACCTCGACATCGGCGACGACTTCCTGGAGATGCCTGGCTGCGACTTTGGC
CGGGGCATGGATGATTTCCTGCAAACCGTCGATTTCTGCGTG**TGA**ATTAAGAGAGGTTTAAGGTTTATCT
TTAAGTAATGCTGGTGTGATGGGCCTCCTGTGAGTGTAATCCGTGTGCATTGGGAGGGAGGATATCCTCT
TGGGTTTTTCCCCGGGCAGCTCAATCAGCTCGTGCCCAAGATGGTGTCTGTTTTCCCTAGTCTCCCAAGC
ACATTAGCGTTATGGTGGTGAGAGTGAGAGAGTATTTTTGATAGTAGGTTACATTGTGATAAGTTATAGT
AGCTGCTGGTTATGTTATGTTCCTCTTTTATGTTACCTGTCAATGTTGGTCCA

>**TaERF1-5** clone C76R1L1-66 1483bp CDS complete, mRNA 137.. 1190 **B-3**

CCCGGTGATTCCGTCCGTCGAATTTTTGACACCCAAGAGAGAGAAAAAGTTGTCCCCTTTTCCTAGTAGA
GTTTTTCCCCTTCAGTTTTGTTAGTCTCAGTCAGTGTTAGAGGAAGAAGAAGAAGAGGCAGGAGCC**ATG**G
CGCCGTTGAGCCAGCACCACATGCGCATCAGGGACGCTCTGGCCAAGAAGACCAAGACCAAGATGCTCGC
CGGGTTTGGCGTCAAGCCCTCGGCTGCTCTCTCCAAGCCGGCGCAGCTGCGGGCGCAGACGCTGCTGCCC
GCTGCGCAGCCACCGAGGCGGCGCGTGCGCGTCCTGTACGAGGACCCCGACGCCACCGACTCGGACACCG
ACGACGAGGAGGCCGCCGCCGCCCCTGCTTCGTCCAAGCGCTGCTTCGAGCTGTTCCTTGGCAAGGCCCC
GCCGGCCAAGTTGTTTGCCAAGCCGGTCACTCCGACCGCCGCGGCTTCTGCCGTCGCTGCCTGCACCACC
AGCAGCGCCGAGGGCCACCGTGGGGTGCGCCTCCGCAAGTGGGGCAAATGGGCGGCCGAGATCCGCAACC
CGTTCTCCGGCAAGAGGGAGTGGCTTGGCACCTTCGACACTGCCGACTTGGCCTCCGCTGCCTACCAGGC
TGCCTCCCGGAGCTTTATCGAGGAGAAGCGTCGCCGCCGTGGCCAGTCTGTGGCCGCGGCCTCACCTGCT
CAGTCGGCTGCGTCAACAACACCGACGTCGTCTTCTACGCCGACGGCATCTTCGTCGTCCTCCACCTCGG
CTGCTCCATTCGCCCACCCGTCGCCGTCCTCTGTCCTGGAAGCTACCAAGCCAGCAGCTGAGTCCCTGTC
GCCTGAGCCGACTCCAGTCCCTGTGATGGTGTCCACACCCGTGGTGTCCACAACGGCTGAGACTGCGCAG
CTTCCTGACGACCCAGAGTTCTACCAGGACCTTCTGCGTGGTCTTCAGCTGCCGGACATTGACCCAATGG
ACTTCCGCGCCGGGCTTGATGCCTTTGGATGTCTCCGAGGCGTCGGCTTACTTGGACGGCGAACAGGACC
TGCTGCTCGGTGACCTGGCCGACGAGGACCTTGAGCTGGACATGGACCTCGACATCGGCGACGACTTCCT
CGACATGCCTGGCTGCGACTTTGGCCGAGGCATGGATGATTTCCTGCAAACCGTCGATTTCTGCGTG**TGA**
ATTAAGAGAGGTTTAAGGTTTTATCTTTAAGTAATGCTGGTGTGATGGGCCTCCTGTGAGTGTAATCCGT
GTGCATTGGGAGGGAGGAGATCCTCTTGGGTTTTTCCCCGGGCAGCTCAACCAGCTCGTGCCCAAGATGG
TGTCTGTCTTCCCTAGTCTCCCGAGCACATTAGCGTTATGGTGGTGAGAGTGAGTGAGAGAGTATTTCTG
ATAGTAGGTTACATTGTGATAAGTTATAGTAGCTGCTGGTTATGTTATGTTCCTCCTTTATGTTACCTGT
CAATGTTGGTCCA

>**TaERF1-6** clone C115R1L1-72 1486bp CDS complete, mRNA 134.. 1180 **B-3**

CCCGGCGATTCCGTCCGTCGAAATTTTTGACACCCGAGAGAAAAAGCTGTCCCCTTTTCCTAGTACAGTT
TTTCCCCTTTCAGTTTTGCTAGTCAGTCAGTGTTAAGAGGAGGAAGAAGAAGAGGCAGGATCC**ATG**GCCC
CGTTGAGCCAGCACCACATGCGCATCAGGGACGCGCTGGCCAAGAGGCCCAAGACCAGGATGCTCGCCGG
GTTCGGCGTCAAGCCCTCCGCCGCCTTCTCCAAGCCGGCGCAGCTGCAGACGCAGGCGCTGCTGCCCGCC
GCGCAGCCGCCGCCGCGGCGGCGCGTGCGCGTCCTGTACGAGGACCCTGACGCCACCGACTCCGACACCG
ACGACGAGGAGGAGGCGGTCGCCGCCCCTGCTTCGTCCAAGCGCTGCTTCGAGCTGTTCCTTGGCAAGGC
CCCGCCGGCCAAGGTGTTTGCCAAGCCGATCACTCCGACTGCCGCGGCTTCTGCTGTCGCTGCCTGCACC
ACCAGCAGCGCCGAGGGCCACCGCGGCGTGCGCCTCCGCAAGTGGGGCAAGTGGGCGGCGGAGATCCGCA
ACCCGTTCTCCGGCAAGAGGGAGTGGCTTGGCACCTTCGACACTGCCGACTTGGCCTCCGCCGCCTACCA
GGCCGCCTCCCGGAGCTTTATCGAGGAGAAGCGTCGCCGCCGTGGCCAGTCTGTGGCCGCGGCCTCGCCT
GCTCGGTCGGCTGCGTCAACAACGCCGACGTCGTCTTCTACTACGCCGACGGCATCTTCGTCGTCCTCCA
CCTCGGCTGCTCCGTTCGCGCACCCGTCGCCCTCCTCTGTCCTGGAGGCCACCAAGCCAGCAGCCGAGTC
CCTGTCGCCTGAGCCGACTCCTGTTCCGGTCATGGTGTCCACCACGGCTGAGACCGCACAGCTTCCCGAC
GACCCAGAGTTCTACCAGGACCTCCTGCGTGGGCTTCAGCTGCCGGACATCGACCCGATGGACTTCCGCG
CCGGGCTTGATGCCCTGGACGTCTCGGAGGCGTCGGCTTACTTGGACGGCGAACAGGACCTGCTGCTGGG
CGACCTGGCCGACGAGGAGCTGGAGCTGGGCATGGACCTCGACATCGGCGACGACTTCCTGGAGATGCCC
GGCTGCGACTTCGGCCGAGGCATGGATGATTTCCTGCAAACCGTCGATTTCTGCGTG**TGA**ATTCAGAGAG
GTTCAAGGTTTATCTTCAAGTAGTAAAATGCTGTGGTGTGATGGGCCTCCTGTGAGTGTAATCAATCCGT
GTGCGTTGGGAGGGAGGAGACCGTCTTGGGTTTTTCCCCGGGCAGCTCAACCAGCTCGTGCCCAAGATGG
TGTCTGTTTTCCCTAGTCTCCCGAGCACATTAGCGTTATGGTGGTGAGAGTGAGTGAGTGAGAGAGTATT
TTTGATAGTAGGTTACATTTGTGATAAGTTATAGTAGCTGCTGGTTATGTTATGTTCCTCCTTCCATGTT
ATCTGTCAAGTGGTCC

>**TaERF1-7** clone CA652588R1L6-45, 742bp CDS complete (115-672)**B-3** Ckc2

GCCACATACAGCCGAAGGAAGAAGAGAACGTCTGTCATCTGCTTAATTCCTTGTGCCGAGGTAGCTGCCA
ATTCCAAAGACTGTGAGTAAGAGGAGAACGTGGTTAGCTAGGAA**ATG**TGCGGCGGAGCTGTTACAGCCGA
CTTTGTCCCGGCCGGAGCCCGCCGCCCGGATGGCTCCTCCGCCGACGTCCCCGGCTCCAGCCTCACCGTC
ACCGGTGAGGAGGTGACGGAGAAATCGCCGGCGCCGGGGCGGAAGACGGCGTACCGTGGGATCAGGCGCC
GGCCATGGGGCCGCTGGGCTGCGGAGATCCGGGACCCCAGGAAGGGCGCGCGCGTCTGGCTGGGCACCTA
CGCCAGCGCGGAGGAGGCCGCGCGCGCCTACGACGTCGCGGCGCGCGATATCCGCGGGCCGAAGGCCAAG
CTCAACTTCCCACCCGCGGTGGGCGCGCCGCAGGAGGCCGCGGCCGTTGCAGGGGCGGGGGCGCCCAAGA
AGCGTCGCATCGTCGCGGCAGAGGAGAGCTCCGCGTCTTGGTCTCCACTTCCGGCCCCGGCTAGCGGCGG
CGGCGGCGGCACAGACAGCCTGCGGGAGCGCATGTCCGGCCTGGAGGCGTTCCTGGGGCTGGAGGACGGC
AACGTGGAGCCCTGGGAGGCCGTCAATCTCATCATGGAG**TAG**GCGTGTGCCGTCGCCACGTGCGGCAGCT
CTGCGGCGTGCATGGGTTACATTTGCCGATAAATAAGCAGAT

>**TaERF1-8** clone ca653896r1l6-49, 879bp CDS complete 131-685 **B-3**Ckc2

GAATTCCCGGGATATCGTCGACCCACGCGTCCGGCCGAAGGAAGAAGAGAACGTCTGTCATCTGCTTAAT
TCCTCTGTGCCGAGGTAGCTGCCAAAGACTGAGTAAGAGGAGAACGTGGTTAGCTAGGAA**ATG**TGCGGCG
GGGCTGTTATCGCCGACTTCGTCCCGGCCGGGGCCCGCCGCCCGGATGGCTCCTCCACCGACGTCCCCGG
CTCCAGCCTCACCGTCACCGGTGAGGAGGTGACGGAGAAACCGCCGGCGCCGGGGCGGAAGACGGCGTAC
CGTGGGATCAGGCGCCGGCCATGGGGCCGCTGGGCTGCGGAGATCCGGGACCCCAGGAAGGGCGCGCGCG
TCTGGCTGGGCACCTATGCTACCGCGGAGGAGGCCGCCCGCGCCTACGACGTCGCGGCGCGCGATATCCG
CGGGCCGAAGGCCAAGCTCAACTTCCCGCCCGCCGTGGGCGCGCCGCAGGCGGCCGCAGCCGTGGAGGGG
GCGGGGGCGCACAAGAAGCGTCGGATGGTCGCGGCAGAGGAGAGCTCCGCGTCTTGGTCTCCACTTCCGG
CCACGGCCACCGGAGGCGGCGGCACAGAGAGCCTGCGGGAGCGCATGTCCGGGCTGGAGGCGTTCCTGGG
GCTGGAGGACGGCGACGTGGAGCCCTGGGAGGCCGTCGATCTCATCTTGGAG**TAG**GCGTGTGCCGTCGCC
ACGTGCGGCGTCCATGGGGTACATTTGCCGATAAATAAGCAGATTGGTCGAAGTCTTTTGATTAGCTAGA
GTATCGTTCTGCTTTCCTGGGCTTTCGCCAGGCATATCCAGGTCCTTTGTAGTGGAATGCTATGTAATTC
GAGCATGACGATTAGAAGAAATGTGGAGAACTTGCAGCA

>**TaEREBP2-1** clone C5F1L1-A(9) 1061bp CDS complete, mRNA 115.. 972 **B-3**

AAGCTGCTCTGAGGTCGTCACTCAAGTACGTAAGAAACAGGGTAGCGACCGACAGGCCAGGGCACGCAGG
CACGTATACCTTAGTCACGTACACACAGCGAGAAGTGAACAAAG**ATG**TGCCGCGGTGCAATCCTCGCCGA
GCTCATACCGAGCGCGCCGCCCAGGAGCGTCACGGCGGTCCACCTCTTGCCCAAGCGGCAGAGGGTGGAC
GACTTCGAGGCTGCTTTCAAGCGCTTCGACGAGGACTCTGAGGAGGAGGAGGGAGGTGCGCCAGTGCTCG
GGGTCGGCGGCAGTAGGCCGGCGCAGTACAGGGGCGTCCGGCGCCGGCCGTGGGGCAAATGGGCGGCAGA
GATCCGCGACCCCGTCAAGGGCGTCCGGTTCTGGCTCGGCACCTTCCCCTCCGCCGACGACGCCGCGCAC
GCCTACGACGACGCCGCCCGCGACTTCCGCGGCGCCAACGCCAGGCTCAACTTCCCCTCCTCGCCCACAA
CAAGCGCACCCAAGTGCCGTGTAGCCGCGAAACCGACACCGTTCGTTGTTATTGACCTCGACGACGACGA
AGAGGGTGATGCTGGAGCAGCGCATGCCGGCGGGATGAGCTCCGAATCCAGCGGCGCCCTGCCGGACTTC
TCATGGCAGGGCATGTCCGCGTCCGGCGAGGTCATGGCGCAATCTGTCCATGCTGAAGTGGAGTCCAGCC
ACTCTGTCGTCGACCTGGGCAGCGCCAAGAAGCGGCCCCGGATCGAAGCCGACGCGGTTCTGCCGGCAGC
GGCCGACGACTCCGCCGACCAACTGTTAGATCCTTTCGTGTTCGATGACCAATTCGGCTTCTCGGACAGC
AGCTCGTACGAGTGGCTGGATGGCCTGTTCGGCGCCGATGCTGAGAAGATCGACGACAGCCAGCTGGGGC
TCTGGAGCTTTGGCGACCATGACTGTCTCGCCGAGGACAGCGCGCGCGTGCAAGTAGAG**TAG**TACTGTAC
TGTATCATCGAAGACGACAGGATAGGACACTCGGACACACCTTGTTGATAATCTCTCCCTCTTGAAAAAA
AAAAAAAAAAA

>**TaEREBP2-2** clone C110F1L8-7 1200bp CDS complete, mRNA 190.. 981 **B-3**

TCATTGATTACGCCAAGCTCAGAATTAACCCTCACTAAAGGGACTAGTCCTGCAGGTTTAAACGAATTCG
CCCTTCAAGCTGCCCTGAGGTCATCATCACTCAAAGTAAGAAAGAGGGTAGCGACCGACAAGCCAGGGGA
CGCAGGCACGTATACCTTGTTACTTACACACAGCAAGAAGTGAGCGAAG**ATG**TGCGGCGGTGCAATCCTC
GCCGAGCTCATACCGAGCGCGCCGCCCAGGAGCGTCACGGCGGTCCACCTCTTGCCCAAGCGGCCGAGGG
TCGACGACTTCGAGGCTGCTTTCAAGCGCTTCGACGAGGACTCTGCGGAGGAGGAGGTGCTCCAGCGCGC
GGGCTTCGAGTTTGGCGCCCACAGGCAGCCGGCAGCGAGGCGCGGTGGCAGTAGTCCGGCACAGTACAGG

GGCGTCCGGGGCCGGCCGTGGGGCAAATGGGCGGCAGATATCCGCGACCCCGTCAAGGGCGTCCGGGTCT
GGCTCGGCACCTTCCCttTtCGCCGAGGCCGCCGCGCACGCCTACGACGACGCCGCCCGCGACTTCCGCGG
CGCCACGCCAGGCTCAACTTCCCCTCTTCGTCCACGAGCGCACCCAAGCGCCACGTGGCCGCGAAGGCGA
CGCCGTGCCTTGTTATTGACCTCGTCGACGACGAAaGAGGGTGATGCTGGAGCAGCGGATGCCGGCGGGA
TGAGTTCCGAATCAAGCGGCGCCCTGCCGGACTTCTCGTGGCAGGGCATGTCCGCGTCCGACGAGGTCAT
GGCGCAATATGTCCATGCTGAAGTGGAGTCCAGCCAGTCCGTCGTCGACCTGGCCAGCCCCAAGAAGCGG
CCCCGGATCGAAGCCGACGCGGTTCTGCCGGCAGCGTCCGACGACTTCGCCGATGCCGAGAAGATCGACG
ACAGCCAGCTGGGGCTCTGGAGCTTTGGCGACGATGACTGTCTCGCGAGGACAGCGCGCGCGTGCAAG**TA**
GATTAGTACTGTATTCATCAAAGACAAAAGGATACGTACGACACTCGCAAGTGTTTGAGGTTCCCCCACA
CTCTGTTGATGTTTTGCTTTCATCAAGAAATTAAAAACTTGTGGTTGAGTAAAACTTGTGGTTTTGGTTT
GAAAAATAGAAAAAGGCGGCGCCCTGAATTCCCCGTGGGGCCAACTAACGGACCCACTTTTTTACAATAG
TCCTATATTG

>**TaRAP2.6**clone C119R1L1-3 1011bp CDS partiale, mRNA 396.. >1011 **B-4**

CTCCCTCCGCGTTCTTATCAGTACGTTGTCCGCGCCTAGGCACCAAAGTCCAAAGCAACAGCCATAGCTC
GATCTCGATCCCCGGCGCGACGAAAGAGAAAGAAGCGGCGGCAGGTCGACAGGTCGATCAACTAAGGTGG
ATCCCCGGAGGCATGGGAAGAGGCCCCTACCCGCCGACGAGGAGGAGGAACAGCCGCCACCGCCGCCGTC
AGCAGCCAAGCACGAGCAGGTGGAGGAGCAGCCGTATCACCCCCTCATCGCGCGCGCTCTGCAGCAGCAA
GGAGCTGCCAGCGCCGGCGGAAGCTCGGGAGCAGATGTGGCCGACCCTTTCCCCGTCACCGGAGGCGTAC
GCGCAGTACTACTACTCGGCGCGCGCCGACCACGACGCCACCGCC**ATG**GTCTCCGCTCTGTCCCACGTCA
TCCGCGCCACACCGGAACAGCAACAAGCCTACTACCCCGCCGGATCCGCCGCTGTCTCAGGAGAACAGCA
GCATCAGCACGATGCGGCGGCTGCCGCGGCCATCGCTGAGGAACAAGGGAGGAAGCGGCACTACAGAGGG
GTGAGGCAGCGGCCATGGGGAAAGTGGGCGGCGGAGATCCGGGACCCCAAGAAAGCGGCTCGTGTATGGC
TCGGCACCTTTGACACGGCTGAGGACGCCGCCATCGCCTACGACGAAGCGGCGCTGCGCTTCAAGGGCAC
CAAGGCCAAGCTCAACTTCCCCGAGCGCGTCCAGGGACGCACCGACCTCGGCTTCGTCGTCACGCGCGGC
ATTCCCGACAGATTGCAGCAACAACAACACTACCCCGCCACCGTGGGGGCGCCGGCAATGCGGCCACCGC
CGCACCAGACCGTGGTGCCGTACCCTGACCTCATGCGGTATGCACAGCTGTTGCAGGGCGCTGGCAGTGC
CGGGGGCGCTGTCAACCTGCCGTTTGGCGCCATGTCGCCCCCGTCGATGTCCTCGTCGTCGCCGCACATA
CTCGACTTCTCGACACAGCAGCTCATCCGAG

>**TaWR-1** clone c24f1l1-a 1117bp CDS complete, mRNA 63. 860 **B-5**

CACCCTCCAGCACCTCTCAACCTCAAGTCAGCGCCACCACCAACAACACCACGAGAACACAC**ATG**ACCTT
CAGCGTCTCGCCGGCGATGGCGGGGGGGCAGGGGCACGAGTACATGATCCGCTTCCACAGCCACTTCGAC
GACCCGTCGCCGAGCACCGCCACCGCCGAGCCGCCGCCGTTTGCCGGAAGGGCGATCTCGCCGGAGCAGG
AGCACGGCGCCATGGTCGCCGCGCTGCTGCACGTCATCTCCGGGTACACCACGCCGCCGCCGGACTTCTT
CTTCCCGGCCGCCGCGAGCAAGGAGGTGTGCCCGGTGTGCAGGGTCGACGGCTGCCTCGGCTGCGAGTTC
TTCGGCGCGGCGGAGGCCACCGGGGCGACCGCAGCGGCATTGGACGCGCCGAAGTCGGCACCGGCGGCGG
TGACCGCGGGCGGGCCGCAGCGGAGGCGGAGGAATAAGAAGAACAAGTACCGGGGCGTCAGGCAGCGGCC
GTGGGGCAAGTGGGCGGCGGAGATCCGCGACCCGCGCCGCGCCGTGCGGGTGTGGCTCGGGACCTTCGAC
ACCGCCGAGGACGCCGCCAGGGCCTACGACCGCGCCGCCGCCGTCGAGTTCCGCGGCCCGCGCGCCAAGCTCA
ACTTCCCCTTCCCCGAGCAGCAGCAGCAGCAGCAGTTGGGCGATGGCAATGCCGCCCCGGCCAAGTCCGA
CACGTGCTCGCCCTCGCCCAGCAGCGCGGAAGTCGATGTCCGGGTCCCGCGGAACGGCGGGCAGGAAACA
GGGGATCAGTTCTGGGACGGCCTGCAGGACCTGATGAAGCTAGACGAGAGCGAGCTCTGGTTCCCGCCAT
CTGGAAATTCTTGGGAC**TGA**ACTGAACCTGCTTGATTAGATCCTAGCCGTTGGAGTGAGTGGACAAGACC
ATTTCACTTTTTTTCTCTTCCTTTTTACCTCTGTTGCATTATTTGGACAAAACAGAGTCTGTGAATTATA
AATAGGATTGTATGAAGATTGAAGAGCCCTCTAGTGAGCAGGGCTGTAAGGTGGGCAAAAAACAGTAAAA
AGCGCGCGTCGGTGATAGAGAAACCCTGCGAGGTGGCAAGAAGGGCGAACTCCAGCTCTCTAACAAA

>**TaWR-2** clone c24f1L8-a, 1004bp CDS complete 63-797 **B-5**

CACGCTCCAGCACCTCTCAACCTCAAGTCAGCGCCACCACCAACAACACCACGAGAACACAC**ATG**ACCTT
CAGCGTCTCGCCGGCGATGGCGGGGGGGCAGGGGCACGAGTACATGATCCGCTTCCACAGCCACTTCGAC
GACCCGTCGCCGAGCACCGCCACCGCCGAGCCGCCGCCGTTTGCCGGAAGGGCGATCTCGCCGGAGCAGG
AGCACGGCGCCATGGTCGCCGCGCTGCTGCACGTCATCTCCGGGTACACCACGCCGCCGCCGGACTTCTT
CTTCCCGGCCGCCGCGAGCAAGGAGGTGTGCCCGGTGTGCAGGGTCGACGGCTGCCTCGGCTGCGAGTTC
TTCGGCGCGGCGGAGGCCACCGGGGCGACCGCAGCGGCATTGGACGCGCCGAAGTCGGCACCGGCGGCGG

TGACCGCGGGCGGGCCGCAGCGGAGGCGGAGGAATAAGAAGAACAAGTACCGGGGCGTCAGGCAGCGGCC
GTGGGGCAAGTGGGCGGCGGAGATCCGCGACCCGCGCCGCGCCGTGCGGGTGTGGCTCGGGACCTTCGAC
ACCGCCGAGGACGCCGCCAGGGCCTACGACCGCGCCGCCGTCGAGTTCCGCGGCCCGCGCGCCAAGCTCA
ACTTCCCCTTCCCCGAGCAGCAGCAGCGCGGAAGTCGATGTCCGGGTCCCGCGGAACGGCGGGCAGGAAA
CAGGGGATCAGCTCTGGGACGGCCTGCAGGACCTGATGAAGCTAGACGAGAGCGAGCTCTGGTTCCCGCC
ATCTGGAAATTCTTGGGACTGAAC**TGA**ACCTGCTTGATTAGATCCTAGCCGTTGGAGTGAGTGGACAAGA
CCATTTCACTTTTTTTCTCTTCCTTTTTACCTCTGTTGCATTATTTGGACAAAACAGAGTCTGTGAATTA
TAAATAGGATTGTATGAAGATTGAAGAGTATACATATTAACAAGGCAGTTTTGGGTTTGTGCTTCAAAAA
AAAAAAAAAAAAAAAAAAAAAAA

>**TaSHINE1-1** clone CA711388F1L1-19 786bp CDS partiale, mRNA 119.. >797 **B-6**

AGGCCGGGCTCACACTCACTCACTCACTCGCACTGCCACACTGCCTGCTGATCATTCCCTTCCCTCTGCC
CGTCCTCTTCCCTGCCAGAGAATCTCCGCTCTGCATTCTGCAGAGAAAC**ATG**GTACAGTCCAAGAAGAAG
TTTCGCGGCGTCAGGCAGCGCCACTGGGGCTCCTGGGTCTCCGAGATTAGGCACCCTCTCCTGAAGAGGA
GGGTGTGGCTGGGCACCTTTGAGACGGCGGAGGAGGCGGCGCGGGCGTATGACGAGGCTGCCATCCTGAT
GAGCGGGCGCAACGCCAAGACCAACTTCCCAGTGCCAAGGAGCGCCAACGGGGAGATCATCGTCGCCCCG
GCAGCGGCACGGGACGGCCGTGGTGTCGGCTCGTCGTCCTCCGGCGCGGCCGGCGCCAGCAGCCTGTCAC
AGATCCTCAGCGCCAAGCTCCGCAAGTGCTGTAAGACACCATCCCCGTCCCTCACCTGCCTCCGCCTAGA
CACCGAGAAGACCCACATTGGCGTCTGGCAGAAGCGCGCGGGAGCCCGCGCCGACTCCAGCTGGGTCATG
ACCGTCGAGCTCAACAAGGAGCCGGCGGCGGTTGCGGCACCAACGCACAGTGACAGCACGGTGTCGGCGA
CTCCTTCCTCGTCGACGTCCACGTCCACAACGGGCTCGCCGCCGGAGACAATGGAGGACGAGGAGAGGAT
CGCGCTGCAGATGATCGAGGAGCTGCTGAGCAGGAGCAGCCCGGCCTCGCCGTCACACTGGCTGCTGCAC
GGTGAAGAGGGCCCTC

>**TaSHINE1-2** clone WEF35N03 1087bp CDS complete, mRNA 126.. 809 **B-6**

AGGCCGGGCTCACACTCACTCACTCGCACTGCCACACTGCCCGCCCGCTGATCATTCCCTCCCCTCTGCC
CGTCCTCCTTCGTGCCAGAGAATCTCCCCGCTCTGCCATTCCGCAGAGAAACATC**ATG**GTACAGTCCAAG
AAGAAGTTTCGCGGCGTCAGGCAGCGCCACTGGGGCTCCTGGGTCTCCGAGATCAGGCACCCTCTCCTGA
AGAGGAGGGTGTGGCTGGGCACCTTTGAGACGGCGGAGGAGGCGGCGCGGGCGTACGACGAGGCCGCCAT
CCTGATGAGCGGGCGCAACGCCAAGACCAACTTCCCCGTGCCGAGGAGCGCCAACGGGGAGATCATCGTC
GCCCCGGCAGTGGCGCGGGACGGCCGCGGTGGCGTCGGCTCGTCGTCCTCCGGCGCGGCCGGCGCCAGCA
GCTTTTCGCAGATCCTCAGCGCCAAGTCCGCAAGTGCTGCAAGACGCCGTTCCCCGTCCCTCACCTGCCT
CCGCCTCGACACCGAGAAGTCCCACATTGGCGTCTGGCAGAAGCGCGCGGGCGCCCGCGCCGACTCCAGC
TGGGTCATGACCGTCGAGCTCAACAAGGAGCCGGCGACAGCGGCGGCGGCACCAACGCCCAGCGACAGCA
CGGTGTCGGCGACTCCTTGCTCGTCCACGTCCACGTCCACAACGGGCTCGCCGCCGGAGGCAATGGAGGA
CGAGGAGAGGATCGCGCTGCAGATGATCGAGGAGCTGCTGAGCAGGAGCAGCCCGGCCTCGCCGTCACAT
GGGCTGCTGCACGGTGAAGAAGGCAGCCTCGTCATC**TGA**AGAAGAAAAAATATTGCACGGTTAAGAAAGT
GTGATCAGGTCACCATCCCAGATCAAGGATCTGGTAGGGTGGTTGGCGCACAGGCAGTTAAGATCATGCA
TTGCTCCACGTCGTAGGTACCAGCTGAGCATCTCCATTACGCACTACGTAAAATCAAGCTTAGGAAACGA
TTAATCACTACCGCGTGTGTGAAGCCCCGTGTATTTATAAATTAATCAAAGGCTTACTTGTATGTAACTA
TGTATATGCCGTTACCGTCAAAAAAAAAAAAAAAAAA

ARTICLE II

THE CBF GENE FAMILY IN HEXAPLOID WHEAT AND ITS RELATIONSHIP TO THE PHYLOGENETIC COMPLEXITY OF CEREAL CBFS

Mohamed Badawi[1], Jean Danyluk[1,2], Barbara Boucho, Mario Houde and Fathey Sarhan

Département des Sciences biologiques, Université du Québec à Montréal, Case Postale 8888, Succursale Centre-ville, Montréal, Québec, Canada, H3C 3P8

[1] These authors contributed equally to this work.

[2] To whom correspondence should be addressed.

Article published in the journal "Molecular Genetics and Genomics 2007 May; 277(5):533-54"

Contribution

I took part in all the stages of experimentation and drafting of the article under the supervision and the support of my director FS and the assistance of JD. I carried out the identification, the cloning and the sequence analyses of the CBF genes. I carried out the studies of expression and chromosomal localization of CBFs. I was involved in the phylogenetic studies. I participated in the sampling of the plant materials. All the authors took part in the drafting of the article.

Résumé

La plupart des plantes en zone tempérée tolèrent le refroidissement des températures et le gel tandis que beaucoup d'espèces des régions tropicales subissent des dommages lorsqu' exposées à des temperatures près de zéro dégré célcius. Pour assurer une bonne survie lors des périodes d'hiver, les plantes mettent en place un mécanisme d'adaptation appelé acclimatation au froid. Ce mécanisme induit l'expression de nombreux gènes nécessaires pour protéger contre les effets causés par les basses températures et le gel. Cette induction est en partie sous le contrôle des facteurs de transcription CBF de la famille AP2/EREBP. Pour comprendre l'évolution et la fonction des facteurs CBF chez les céréales, 15 différents gènes CBF de blé hexaploid ont été clonés, identifiés et caractérisés. Cependant, une analyse phylogénétique indiquent que les espèces de blé (*T. aestivum* et *T. monococcum*) peuvent contenir jusqu'à 25 gènes CBF différents. Une classification de tous les CBFs chez les *Poaceae* a permis de distinguer dix groupes qui partagent une origine phylogénétique commune et des caractéristiques structurales semblables. Six de ces groupes (IIIc, IIId, IVa, IVb, IVc et IVd) se retrouvent exclusivement chez les espèces *Pooideae* particulièrement très tolérantes aux basses températures et au gel. Ce résultat indique que les membres CBF de ces six groupes seraient impliqués dans les réponses adaptatives récemment apparues durant la colonisation des habitats tempérés. Des études d'expression ont indiqué que cinq des groupes spécifiques aux *Pooideae* accumulent leur transcrits durant une exposition au froid et de façon constitutive. Cette accumulation de transcrits est plus élevée chez un cultivar d'hiver que chez un cultivar de printemps. Ce profil d'expression semble donc refléter un caractére acquits ayant un rôle prédominant dans la capacité de tolérance aux basses températures chez les cultivars d'hiver, des ces espèces de *Pooideae*.

Mots clés: blé, *Pooideae*, CBF/DREB, AP2/ERF, L'évolution, à basse température

Abstract

Most temperate plants tolerate both chilling and freezing temperatures whereas many species from tropical regions suffer chilling injury when exposed to temperatures slightly above freezing. Cold acclimation induces the expression of cold-regulated genes needed to protect plants against freezing stress. This induction is mediated, in part, by the CBF transcription factor family. To understand the evolution and function of this family in cereals, we identified and characterized 15 different *CBF* genes from hexaploid wheat. Our analyses reveal that wheat species, *Triticum aestivum* and *T. monococcum,* may contain up to 25 different *CBF* genes, and that *Poaceae CBF*s can be classified into ten groups that share a common phylogenetic origin and similar structural characteristics. Six of these groups (IIIc, IIId, IVa, IVb, IVc and IVd) are found only in the *Pooideae* suggesting they represent the CBF response machinery that evolved recently during colonization of temperate habitats. Expression studies reveal that five of the *Pooideae*-specific groups display higher constitutive and low temperature inducible expression in the winter cultivar, and a diurnal regulation pattern during growth at warm temperature. The higher constitutive and inducible expression within these CBF groups is an inherited trait that may play a predominant role in the superior low temperature tolerance capacity of winter cultivars, and possibly be a basis of genetic variability in freezing tolerance within the *Pooideae* subfamily.

Key words: Wheat, *Pooideae*, CBF/DREB, AP2/ERF, evolution, low temperature

Introduction

The ability of plants to respond and adapt to low temperature (LT) stresses varies greatly with species (Sakai and Larcher 1987). Most temperate plants, such as wheat, canola and *Arabidopsis*, are able to tolerate both chilling and freezing temperatures. In contrast, species from tropical regions, such as tomato, maize and rice are unable to tolerate freezing temperatures and even suffer chilling injury when exposed to temperatures in the range of 0 to 12°C. The cold acclimation process induces the expression of cold-regulated (*COR*) genes, whose products are thought to be necessary for protection against freezing stress (Thomashow 1999). Many of these *COR* genes contain copies of the C-repeat/dehydration-responsive element (CRT/DRE) in their promoters, which has the core motif CCGAC and is responsible for the LT-responsiveness of these genes. The factors that bind the CRT/DRE were first identified in *Arabidopsis* and designated CRT-Binding Factors/DRE-binding proteins 1 (*CBF/DREB1*) genes (Stockinger et al. 1997; Liu et al. 1998). These genes encode LT-induced transcription factors that, when constitutively overexpressed in *Arabidopsis*, mimic cold acclimation by inducing *COR* gene expression and freezing tolerance (FT) (Jaglo-Ottosen et al. 1998; Liu et al. 1998). CBF-like proteins have been isolated from a wide range of plants that include species capable and incapable of cold acclimation, suggesting that the CBF cold response pathway is broadly conserved in plants (Jaglo et al. 2001).

The CBF/DREB1 proteins belong to the AP2/ERF superfamily of DNA-binding proteins. This protein family has in common the AP2 DNA-binding motif. The 122 *Arabidopsis* members of the ERF family containing one AP2 DNA-binding motif were divided into 12 groups, with several of the groups being further divided into subgroups (Nakano et al. 2006). The six CBF/DREB1 proteins of *Arabidopsis* were included in subgroup IIIc, one of the five group III subgroups. The CBF proteins are distinguished from other group III members by a conserved set of amino acid sequences (motif CMIII-3) flanking the AP2 DNA-binding domain. Other motifs

found in CBF proteins (CMIII-1, CMIII-2 and CMIII-4) are also present in one or more of the group III subgroups, suggesting related molecular functions may be conserved between subgroups. These features were previously noted as conserved features of the C-terminal activation domain between aligned CBFs (Wang et al. 2005; Skinner et al. 2005). Comparison of group III proteins between the eudicot *Arabidopsis* and the monocot rice reveals they share 4 common subgroups (Nakano et al. 2006) suggesting a functional diversification of group III proteins had occurred before the divergence of these two species. The four ancestral genes have amplified to 23 and 26 genes in *Arabidopsis* and rice, respectively. *Arabidopsis* CBF studies have demonstrated that *CBF1*, *2*, and *3* function in the cold-acclimation pathway with redundant and some possibly specific functions (Gilmour et al. 2004; Novillo et al. 2004; Van Buskirk and Thomashow 2006), *CBF4* is involved in drought adaptation (Haake et al. 2002), *DDF1* and *DDF2* are involved in gibberellin biosynthesis and salt stress tolerance (Magome et al. 2004). Based on these findings, the functional divergence of group III ERF proteins in monocots may differ from eudicots and therefore warrant a characterization of monocot genes.

Members of the *Poaceae* have been targeted for study since they contain the major cereal crops wheat, maize and rice which provide >60% of the calories and proteins for our daily life. To meet the needs of the projected human population by 2050, cereal grain production must increase at an annual rate of 2% on an area of land that will not increase much beyond the present level (Gill et al. 2004). Therefore, significant advances in our understanding of the *CBF* family in cereals are essential to develop needed strategies to protect crops from losses caused by abiotic stress. The *Poaceae* represent an excellent model system to study the roles of the *CBF* family in the evolution of LT tolerance. The *Poaceae* radiated some 55-70 million years ago (MYA) into several subfamilies (Kellogg et al. 2001). The subfamilies *Oryzaceae* (rice) and *Panicoideae* (maize) have a more tropical geographical distribution compared to members of the *Pooideae*, which contain the temperate cereals wheat, barley and oat. LT tolerance within the *Pooideae* subfamily ranges from low in oat (a *Poeae* tribe representative), to intermediary in barley and wheat (*Triticeae* tribe), to

highly tolerant in rye (*Triticeae* tribe). The estimated divergence time between the *Triticeae* and *Poeae* is around 35 MYA, and within the *Triticeae*, barley and rye diverged from wheat around 11 and 7 MYA, respectively (Huang et al. 2002b). The more recent evolutionary history of bread wheat started with an adaptive radiation of the diploid progenitors around 2.2-4.5 million years ago followed by successive hybridizations around <0.5 MYA and 8000 years ago to produce hexaploid bread wheat (Huang et al. 2002c). Wheat is an interesting model since the comparative analysis of *CBF* gene function among closer and more distantly related species may shed light on important evolutionary trends that have sculpted *CBF* function. Furthermore, the 3 genomes of hexaploid wheat are known to contain differences for many agronomically important genes (Gill et al. 2004), and recently, a LT tolerance QTL on chromosome 5 of *T. monococcum* (Miller et al. 2006), barley (Skinner et al. 2006) and hexaploid wheat (Båga et al. 2006) was found to coincide with the location of 11, 12 and 2 *CBF* genes, respectively. The exact molecular explanation for this LT tolerance QTL is not known but indicates the possibility that *CBF* genes may be at the base of this important trait.

Many *CBF* genes have been identified from *Poaceae* species such as rye (Jaglo et al. 2001), rice (Dubouzet et al. 2003; Skinner et al. 2005), barley (Choi et al. 2002; Xue 2002; Francia et al. 2004; Skinner et al. 2005), wheat (Jaglo et al. 2001; Kume et al. 2005; Skinner et al. 2005; Vágújfalvi et al. 2005; Miller et al. 2006), and *Festuca arundinacea* (Tang et al 2005). These studies, in particular those of Skinner et al. (2005) and Miller et al. (2006), have revealed that the cereal *CBF* family is large and complex. To better understand the functions of this gene family during cold stress and its evolution in the *Poaceae*, we initiated a study to identify and characterize *CBF* genes from hexaploid wheat. Here, we show that hexaploid wheat contains at least 15 different *CBF* genes, and that *Poaceae CBF*s can be subdivided into groups with specific characteristics. These findings expand our understanding on the functional categories of cereal *CBF* genes and provide a starting point for future studies.

Materials and methods

Preparation of the cDNA libraries

Five different cDNA libraries prepared from *Triticum aestivum* L. cv Norstar were used to identify expressed wheat *CBF* genes in this study (Table S1). Plant growth conditions, RNA purification and cDNA library construction were described in detail elsewhere (Houde et al. 2006). Briefly the five libraries (L2-L6) were prepared from the following pooled mRNA populations: L2) Aerial parts (leaf and crown) from control and long-term cold acclimated wheat (1 to 53 days); L3) Root tissue from control, cold-acclimated and salt stressed wheat; L4) Aerial parts of dehydration stressed wheat; L5) Crown tissue during vernalization and different developmental stages of spike and seed formation in wheat; L6) Crown and leaf tissues from wheat after short times exposures to LT in the light and in the dark. All cDNAs synthesized were directionally cloned into the pCMV.SPORT6 vector (Invitrogen) with the *Sal*I adaptor (GTCGACCCACGCGTCCG) and *Not*I primer adaptor ($GCGGCCGCCCT_{15}$). For the last 4 libraries, the first strand cDNA reaction mix contained methylated dCTP to prevent cDNAs from internal cleavage by the *Not*I restriction enzyme used for directional cloning. For the last library, the 'GeneRacer' kit (Invitrogen) was used prior to first strand synthesis to produce a library containing a high proportion (95%) of full length cDNAs. For each library, six million primary transformants were obtained, amplified and frozen as glycerol stocks.

To prepare plasmids for PCR experiments, 40µl of a bacterial library stock (>40 x 10^6 clones) were inoculated into LB media (100ml) supplemented with ampicillin and grown at 30°C for 6-8 hours. Plasmids were isolated using the QIAprep Miniprep system (QIAGEN) and the quantity was estimated on a gel.

Identification of wheat *CBF* genes

The gene cloning approach, gene names and GenBank accession numbers are summarized in Table 1. To initiate this project, available CBF protein sequences were

used for data mining of the NCBI NR and EST databases in search of wheat homologs. The mRNA (Jaglo et al. 2001) and EST sequences were assembled into virtual mRNAs using CAP3 (http://pbil.univ-lyon1.fr/cap3.php) (Huang and Madan 1999). This initial assembly was updated as additional *CBF* genes were sequenced from the project. Virtual mRNAs that contained EST sequences from the Wheat Genomics of Abiotic Stress (WGAS) project (https://www.bioinfouqam.wgas.ca/cgi-bin/abiotic/project.cgi) were ordered and completely sequenced at the McGill University and Genome Quebec Innovation Center (Montreal, Canada). Initially, hybridization probes corresponding to either incomplete *CBF* genes or to their AP2 DNA binding region were used to screen the wheat plasmid cDNA libraries L3 and L6 to identify additional full length *CBF* clones. Because this approach was time consuming, a PCR strategy was used to identify *CBF* genes expressed in wheat. From one to three gene specific primers (GSPs) were designed using Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) (Rozen and Skaletsky 2000) in the forward (5' mRNA region) or reverse (3' mRNA region) directions on the assembled virtual mRNAs. In order to drive the PCR reaction, the GSPs were designed with a higher Tm than the universal primers M13F 5'-AGATCCCAAGCTAGCAGTTTTCCCAGTCACGA-3' and M13R 5'-GAGCGGA TAACAATTTCACACAGGAAACAGCTATGA-3' found in the pCMV.SPORT6 vector. A list of GSPs used in the isolation of *CBF* genes is given in Table S2. Amplification was performed using a GSP (0.6 μM) and the corresponding universal primer (0.3 μM), 10 ng of plasmid DNA from a library, 1X or 2X enhancer solution, and *Pfx* DNA polymerase (Invitrogen) following the manufacturer's recommendations. Briefly, the PCR thermal-cycling parameters were initiated with a progressive step down in annealing temperature that ended with 45 cycles of 94°C for 20 s, 64°C for 20 s and 68°C for 150 s. PCR products were analyzed on a gel and lanes that produced DNA fragments of expected size were chosen for subcloning using the Zero Blunt TOPO PCR Cloning Kit (Invitrogen) and then sequenced. Novel CBF sequences were fully sequenced and overlapping sequences were merged to produce the longest possible gene sequence (Table 1). Several clones from

independent amplification reactions were sequenced in order to correct errors introduced during PCR.

Expression profiling using northern analysis

The spring wheat cultivar *Triticum aestivum* L. cv Quantum and the winter cultivar Norstar were germinated in moist sterilized vermiculite in a growth chamber (Model-E15, Conviron) for 7 days at 20°C and 70% relative humidity under an irradiance of 250 μmol m^{-2} sec^{-1} and a 16-h photoperiod. At the end of this period, 5 g of control leaves were sampled and individually frozen. A cold treatment (4°C) was initiated by changing the temperature in the growth chamber 4 hours into the day phase and continued under the same irradiance and photoperiod conditions for the times indicated in Figure 4. Total RNA was extracted from wheat leaves as described (Danyluk and Sarhan 1990), and equal amounts (10μg) were separated on formaldehyde-agarose gels. Transfers to positively charged nylon membranes and hybridizations with ^{32}P-labeled probes were performed using standard molecular biology techniques. Probes for the different *TaCBFs* were designed outside of the AP2 domain to avoid cross-hybridization with known CBF transcripts (Table S3). All filters were washed at high stringency (0.1X SSC, 0.1% SDS; 65°C) for 30 minutes. Membranes were exposed to BioMax-MS films (Kodak) and K screen (Bio-Rad).

Expression profiling by quantitative real-time PCR

Plant growth conditions and cDNA synthesis

Spring cultivar Manitou and winter cultivar Norstar were germinated in a mixture of 50% black earth and 50% Pro-Mix (Premier) for 8 days at 20°C under a 16-h photoperiod with a light intensity of 250 μmol m^{-2} sec^{-1}. In this experiment, Manitou was used as the spring cultivar in order to compare the response to the Quantum spring cultivar. At the end of this period, 8 control seedlings were sampled every two hours on dry ice, and immediately frozen at -70°C. Cold treatment (4°C) was initiated by changing the temperature in the growth chamber 4 hours before the night cycle, and sampling as indicated in each Figure. Sampling of control and cold-

treated plants every two hours was chosen so that closely spaced time points act as closely related biological replicates in addition to providing a better resolution of the expression characteristics of a gene. Night samples were collected by opening the growth chamber in the dark, removing plants to another room and harvesting rapidly in the light. Total RNA was isolated using the RNeasy Plant Mini Kit (QIAGEN) using the optional on-column DNAse digestion. Purified RNA (2.8 μg) was reverse transcribed in a 20 μL reaction volume using the Superscript II first strand cDNA synthesis system for RT-PCR (Invitrogen). Parallel reactions for each RNA sample were run in the absence of Superscript II (no RT control) to assess for genomic DNA contamination. The reaction was terminated by heat inactivation; the cDNA product was treated with RNase H, and diluted in water (20 ng/μl) for storage (–20°C).

Design of gene-specific primers

The genome of hexaploid wheat contains three genomes inherited from three diploid ancestors. The 37 *TaCBF* gene sequences identified in this study were analyzed using ClustalW (http://align.genome.jp/) and phylogenetic characterization. This analysis revealed 15 groups of genes containing one to three homeologous copies in each group. Primers were specifically designed to monitor the expression of only one representative gene per group. This representative was chosen randomly. Fluorescent TaqMan-MGB probes as well as the non-fluorescent primers (Table S4) were designed using the combination of Primer Express Software Version 2.0 (Applied Biosystems) and Primer3. BLASTN searches against EST and NR databases were performed to confirm the gene specificity of the primers. Non-fluorescent primers were synthesized by Invitrogen and TaqMan-MGB probes by Applied Biosystems.

PCR amplification and data analysis

Quantitative real-time PCR assays for each gene target were performed in quadruplicate on an ABI Prism 7000 sequence detection system (Applied Biosystems) using the eukaryotic 18S rRNA as the endogenous control (Applied Biosystems #

4319413E). From the diluted cDNA, 2 µL (40 ng) was used as a template in a 25-µL PCR reaction containing 1X TaqMan universal PCR master mix (Invitrogen), 0.9 µM of non-fluorescent primers, and 0.25 µM of TaqMan-MGB fluorescent probe. The PCR thermal cycling parameters were 50°C for 2 min, 95°C for 10 min followed by 50 cycles of 95°C for 15 s and 60°C for 1 min.

All calculations and statistical analysis were performed by the SDS RQ Manager 1.1 software using the $2^{-\Delta\Delta Ct}$ method with a relative quantification $(RQ)_{min}/RQ_{max}$ confidence set at 95 % (Livak and Schmittgen 2001). The error bars display the calculated maximum (RQ_{max}) and minimum (RQ_{min}) expression levels that represent SE of the mean expression level (RQ value). Collectively, the upper and lower limits define the region of expression within which the true expression level value is likely to occur (SDS RQ Manager 1.1 software user manual; Applied Biosystems). Amplification efficiency (90% to 100%) for the 15 primer sets was determined by amplification of cDNA dilution series using 80, 20, 10, 5, 2.5, and 1.25 ng per reaction (data not shown). Specificity of the RT-PCR products was assessed by gel electrophoresis.

Chromosome localization of TaCBF genes

Genomic DNA was extracted and quantified (Limin et al. 1997) from several stocks of the wheat cultivar Chinese Spring: ditelocentric series provided by the USDA from E. R. Sears collection (all chromosomes are present but in each line one chromosome pair is represented by only the telocentric chromosomes of one arm); chromosome 5 nullisomic-tetrasomic lines (a pair of chromosomes is removed and replaced by another pair of homoeologous chromosomes); and deletion lines for homoeologous group 5AL and 5DL chromosomes (Endo 1988; Endo and Gill 1996) generated using the gametocidal chromosome of *Aegilops cylindrical*. From the diluted genomic stocks, 2 µL (20 ng) was used as a template in a 25-µL PCR reaction containing 1X TaqMan universal PCR master mix (Invitrogen), 0.9 µM non-fluorescent primers, and 0.25 µM TaqMan-MGB fluorescent probe. The PCR thermal cycling parameters were 50°C for 2 min, 95°C for 10 min followed by 50 cycles of

95°C for 15 s and 60°C for 1 min. At the end of the run, the Ct values were compared, and genetic stocks that showed a delayed or undetectable amplification were identified as the location of the assayed *TaCBF* gene.

Phylogenetic and other bioinformatic analysis

Monocotyledonous *CBF* homologs were identified using the *TaCBF* nucleotide and protein sequences as queries against the GenBank NR and EST databases. Overlapping ESTs were assembled into virtual cDNAs using CAP3 (http://pbil.univ-lyon1.fr/cap3.php) (Huang and Madan 1999) and a consensus cDNA sequence was deduced. EST-derived sequences for analyses were obtained by trimming edges that corresponded to low quality error prone regions which were revealed through blastx searches against the NR database. Accession numbers for genes and ESTs used in this study are in Table 2. FASTA files of nucleotide and protein sequences used in this analysis are presented in Supplemental Tables S5 and S6. The degree of sequence identity was determined using ALIGN and FASTA on the Biology Workbench (http://workbench.sdsc.edu). Sequences were aligned using ClustalW from the Biology Workbench or from the MEGA software version 3.1 (Kumar et al. 2004), and alignments were refined manually. The MEGA software was used for phylogenetic analyses and the Minimum Evolution tree was derived from this alignment using the Kimura 2-parameter with bootstrap test and default parameters.

Hydrophobic cluster analysis (HCA) (Gaboriaud et al. 1987; Callebaut et al. 1997) was conducted using the web-based interface at: (http://bioserv.rpbs.jussieu.fr/RPBS/cgibin/Ressource.cgi?chzn_lg=an&chzn_rsrc=HCA). Briefly, the protein sequences are displayed on a duplicated α-helical net in which hydrophobic amino acids (V, I, L, F, M, Y, W) are contoured. Hydrophobic residues separated by four or more nonhydrophobic residues, or a Proline, are placed into distinct clusters. The defined hydrophobic clusters were shown to mainly correspond to the internal faces of regular secondary structures (α-helices or β-strands). Two other amino acids were chosen to be highlighted in this study: proline

which confers the greatest constraint to the polypeptide chain and glycine which confers the largest freedom to the chain. This secondary structure information was highlighted on a ClustalW alignment of group-related wheat CBFs.

RESULTS

Identification of wheat CBF genes

CBF family members are important regulators of FT in plants. Data mining and analyses of cereal *CBF* sequences present in GenBank suggest that different species contain diverse and complex *CBF* families. Based on preliminary studies conducted in a number of varieties, hexaploid wheat contains at least seven *CBF* genes (Jaglo et al. 2001; Kume et al. 2005; Skinner et al. 2005). To maximize our chance of discovering *CBF* genes involved in the development of wheat FT, several cDNA libraries were constructed (Table S1) and screened to identify *CBF* genes expressed under various cold acclimation time points and conditions in the freezing tolerant cultivar Norstar. A combination of EST sequencing, cDNA library screening and PCR amplification allowed the identification of 37 expressed *TaCBF* genes from hexaploid wheat (Table 1). To be consistent with the established *H. vulgare* and *T. monococcum* nomenclature (Skinner et al. 2005; Miller et al. 2006), we assigned identical gene numbers to orthologs of hexaploid wheat (2 to 15) and new consecutive numbers (from 19 to 22) to novel genes identified following homology comparison with published and identified wheat sequences (e.g. *TaCBFI-A11* shows the highest homology with its ortholog *HvCBFI1* identified previously; Skinner et al. 2005). These analyses revealed that the 37 genes identified from hexaploid wheat can be classified into at least 15 different orthologous gene groups with 1 to 3 homeologous copies in each. Phylogenetic analysis of *T. aestivum* and *T. monococcum* genes reveals that wheat *CBF* genes can be divided into 10 monophyletic groups. Therefore, we included in the proposed CBF nomenclature CBF subgroup information (e.g. CBFIa, II, IIIa, IIIb, IIIc, IIId, IVa, IVb, IVc and IVd) which should facilitate future comparison of monocot CBF properties and functions. In the cases where homeologous copies were mapped to one of the three genomes of hexaploid wheat, a letter designating its location precedes the *CBF* gene number (e.g.

TaCBFIIIa-D6). On the other hand, when the genomic localization of a homeologous copy has not yet been determined, the *CBF* gene number is followed by a temporary designation of .1, .2 or .3 (e.g. *TaCBFIIIa-6.1*).

At least three gene groups (*TaCBFIIId-19*, *TaCBFIVb-20* and *TaCBFIVd-22*) represent true orthologous series in hexaploid wheat since the homeologous copies (A, B and D) were identified and mapped to each genome equivalent in our study (Table 1). One gene group (*TaCBFII-5*) was not mapped in our study (Table 1). However, the *T. monococcum* ortholog was mapped to chromosome 7A (Miller et al. 2006) and the barley ortholog was mapped to the short arm of 7H (Skinner et al. 2006) suggesting that this gene group will be located on chromosome 7. The vast majority of gene groups (13 out of 15) were mapped to chromosome 5 (Table 1), and at least 9 of these 13 have so far been mapped more precisely to a region, between two deletion breakpoints, associated with a cold tolerance QTL in several *Triticeae* species (Francia et al. 2004; Båga et al. 2006; Miller et al. 2006; Skinner et al. 2006). The present study allowed the localization of 4 new gene groups (*TaCBFIIId-19*, *TaCBFIVb-20*, *TaCBFIVb-21* and *TaCBFIVd-22*) (Table 1) to this QTL region containing the 11 tandem *CBF* genes in *T. monococcum* (Miller et al. 2006) and 12 tandem *CBF* genes in barley (Skinner et al. 2006).

A cut-off of 95% was chosen to differentiate between homeologous copies and possible recently duplicated genes. Because homeologous copies of different *CBF* gene groups would have started diverging around the same time, one would expect them to show similarities in the same range. The comparison of *T. monococcum* (Miller et al. 2006) and *T. aestivum CBF* genes (this study) did reveal five orthologous genes with high levels of identity (more than 98%). From the 13 *T. aestivum* gene groups that contain homeologous copies, only 5 show identities below 95% (the percent identity excludes bases involved in transitions, transversions and gaps) between the homeologous ORFs. These include *TaCBFII-5*, *TaCBFIIIa-6*, *TaCBFIIIc-3*, *TaCBFIVd-9* and *TaCBFIVd-22* which show identities of 85.9, 91.7, 94.5, 94.8 and 89.9% respectively. When the identity comparison is repeated with gap regions not included, only *TaCBFII-5* and *TaCBFIIIa-6* show lower similarities

of 93.4% and 93.5%, respectively. Therefore, these results suggest that the *TaCBFII-5* and *TaCBFIIIa-6* homeologous groups may contain closely related paralogs and/or have diverged at a different rate than other groups.

Phylogeny of monocot CBF genes

Monocot and eudicot *CBF* sequences are separated on a phylogenetic tree (Dubouzet et al. 2003; Qin et al. 2004; Bräutigam et al. 2005; Xiong and Fei 2006) suggesting that at least some *CBF* gene function specialization has evolved recently in plants. To understand the relationship between *TaCBF*s and other monocot *CBF*s, we searched NR and EST databases and compiled a set of sequences for analysis (Table 2). During our BLAST comparisons of *CBF*s, we found that it would be difficult to align the entire ORF of distant members with complete confidence. Therefore, the nucleotide sequence of the AP2 DNA binding domain and adjacent CBF signatures were chosen for alignment and phylogeny analysis since these domains are extremely well conserved in the CBF family. To simplify the future comparison of *CBF* gene functional studies with different monocotyledonous plants, we propose a nomenclature that reflects the evolutionary relationship of *CBF* genes. This will help to distinguish their specific functional roles which may have appeared during their evolution.

To establish the relationship between the different *CBF* genes in wheat, only one representative of each of the 15 *TaCBF* gene groups (preferentially the homeologous A copy) was included in this analysis. In addition, 8 (Table 2) of the 13 identified *T. monococcum CBF* genes (Miller et al. 2006) were included in the analysis since they showed less than 95% identity (gap regions not included in the analysis) with any of the 15 *TaCBF* gene groups. These lower identities of *TmCBFII-5* (86.5%), *TmCBFIIIb-18* (76%), *TmCBFIIIc-10* (94%), *TmCBFIIIc-13* (86.7%), *TmCBFIIId-16* (82.1%), *TmCBFIIId-17* (78.8%), *TmCBFIVa-2* (92.4%) and *TmCBFIVd-4* (87.1%) with their closest homologues suggest that they may represent additional wheat *CBF* genes. To compare wheat *CBF*s with the closely related

Triticeae species barley, we included 13 of the 19 *HvCBF* genes reported (Table 2) (Skinner et al. 2005). The two additional pseudogenes, *HvCBFIIIc-8B* and *HvCBFIIIc-8C*, and the remaining 4 genes *HvCBFIIIc-10B* (97.2%), *HvCBFIVa-2B* (98.8%), *HvCBFIVd-4B* (99.8%) and *HvCBFIVd-4D* (96.4%) were not included since their high homologies with their closest related paralog suggested that these duplications happened following the divergence of wheat and barley. Several other *CBF* sequences from families of the monocotyledonous order *Poales* were included in the analysis (Table 2). In addition, *CBF* representatives of two other monocotyledonous orders *Arecales* and *Zingiberales* were also included in the analysis (Table 2). *HvCBF7* and *TmCBF7* (Skinner et al. 2005; Miller et al. 2006) were not included in this phylogenetic analysis since their protein sequence contains a less conserved CBF signature and the presence of motifs found in subgroup IIId of ERF proteins (Nakano et al. 2006).

The phylogenetic analysis presented in Figure 1 shows that monocot *CBF*s cluster into several distinct monophyletic groups. Observation of the first group (named CBFI) reveals 3 distinct branches. The first branch containing the *CBF* sequences from the *Arecales* and *Zingiberales* orders (*SpCBFI*, *RhCBFI*, *DlCBFI* and *ZoCBFI*) is separated from the two branches containing *Poales CBF* sequences. As additional *CBF* genes are identified and characterized in other monocotyledonous plants, it may reveal functional differences that will necessitate a classification as a distinct subgroup to better reflect their evolutionary relationships. The second branch contains the lone rice gene *OsCBFI-1F*. If orthologs of this gene are found in other *Poales* members, this will support the need to define a separate subgroup representing these more distantly clustered genes. The remaining branch, tentatively named CBFIa, contains sequences from the *Poales*/*Poaceae* subfamilies *Oryzaceae*, *Panicoideae* and *Pooideae* suggesting that the ancestral *Poaceae CBFIa* was present before divergence of these subfamilies, and that none of these families have lost this *CBF* gene. In fact, *Oryzaceae* and *Pooideae* may have already possessed 2 genes before divergence since rice and barley each have a pair of genes (*CBFIa-1* and *CBFIa-11*). Although only *TaCBFIa-A11* has been identified to date, the above information

suggests that wheat could have one or two additional genes (orthologs of CBFIa-1 and *OsCBFI-1F*). The characteristic of group CBFI is that it is the only one that contains *CBF* genes from the three orders *Poales*, *Arecales* and *Zingiberales*. Although the identification of different *CBF* genes from orders other than *Poales* is still limited, their clustering with the CBFI group suggests that it is the most ancient group in monocots. In support of this, proteins encoded by *CBFI* genes are the ones that show the highest homologies with dicotyledonous CBF proteins suggestive of their closer evolutionary relationship with an ancestral type CBF. This was also noted previously by Skinner et al. (2005).

The second group (named CBFII) also contains *CBF* genes from the *Oryzaceae*, *Panicoideae* and *Pooideae* subfamilies but the presence of only one gene in rice suggests that it was less complex before divergence. The possibility that additional genes may exist in wheat based on the lower homology of *TmCBFII-5* (86.5%) with *T. aestivum* genes will require additional sequence identification and characterization to be confirmed. Although these genes were initially presented as part of group I (Skinner et al. 2005), they are classified separately here as a group II based on their evolutionary distance from the first group, their specific occurrence in all *Poales/Poaceae* subfamilies examined, and structural differences (see next section).

The results presented in Figure 1 show that the 11 *CBF* genes from group 3 are clustered in several distinct subgroups. These subgroups were named CBFIIIa, CBFIIIb, CBFIIIc and CBFIIId to reflect their monophyletic origins. Groups IIIa and IIIb are the only ones that contain *CBF* genes from the 3 *Poaceae* subfamilies suggesting that the ancestral *CBFIIIa* and *CBFIIIb* genes were already present before divergence of these subfamilies. However, group CBFIIIb did not always form a monophyletic clade with other substitution models suggesting a less certain orthologous relationship. With these alternative models, *TmCBFIIIb-18* was found to cluster alone or in the vicinity of the CBFIIId group (results not shown). Sequencing of additional *CBFIIIb* genes will help to resolve this ambiguity. Interestingly, only *CBF* genes from tribes of the *Pooideae* subfamily were found clustered in groups

CBFIIIc and CBFIIId suggesting that these groups evolved following the appearance of the *Pooideae*. Based on the available data, group IIIc contains at least 3 common genes (*CBFIIIc-3*, *CBFIIIc-10* and *CBFIIIc-13*) before wheat-barley speciation. No report has yet shown the existence of the *HvCBFIIIc-8* type pseudogenes in wheat. In addition, this group contains two of the genes (*HvCBFIIIc-8* and *HvCBFIIIc-10*) that have duplicated in barley following divergence from wheat. Excluding the pseudogenes, this group is presently composed of 4 genes in barley and 4 genes in wheat. In the case of group IIId, one barley gene has been identified (*HvCBFIIId-12*) compared to five in wheat (*TaCBFIIId-12*, *TaCBFIIId-15*, *TmCBFIIId-16*, *TmCBFIIId-17*, *TaCBFIIId-19*). However, it is probable that barley has at least an additional *CBFIIId* gene since multiple homologs were identified in the more distantly related *Avena sativa*. These results suggest that differences may exist between closely related species in the exact number of *CBF* genes present within a group. The group III rice genes *OsCBFIII-1D*, *OsCBFIII-1I* and *OsCBFIII-1J* do not cluster within any of the above described subgroups. Therefore, no subgroup classification is proposed since these genes may have evolved specifically in *Oryzaceae*.

The analysis of the 9 wheat genes from group IV (Figure 1) reveals a compact clustering compared to group III genes, suggesting a more recent diversification. However, based on the origin of the main branches, it is possible to classify the wheat genes into four groups: CBFIVa (*TaCBFIVa-2* and *TmCBFIVa-2*), CBFIVb (*TaCBFIVb-20* and *TaCBFIVb-21*), CBFIVc (*TaCBFIVc-14*), and CBFIVd (*TaCBFIVd-4*, *TmCBFIVd-4*, *TaCBFIVd-9* and *TaCBFIVd-22*). The group CBFIVb is the only one that does not contain a barley representative at the moment. Within group CBFIVd, the identity between the complete ORF of wheat-barley orthologs (91.5% for *CBFIVd-4* and 94.2% for *CBFIVd-9*) and paralogs (90.8% for wheat-wheat and 90.3% for barley-barley) are very similar indicating that this group amplified just prior to the divergence of the wheat-barley lineage. At present, only two genes (*AsCBFIVa* and *FaCBFIVa-2*) from the closely related *Pooideae* tribes *Aveneae* and *Poeae* were found to cluster with the CBFIVa group indicating that at

least this group has appeared prior to the radiation of these tribes. These results suggest that the amplification of group CBFIVd and possibly the emergence of groups IVb and IVc may represent a specific characteristic of the *Triticeae* tribe. In addition, groups CBFIVa and CBFIVd contain the two genes (*HvCBFIVa-2* and *HvCBFIVd-4*) that have duplicated in barley following divergence from wheat. The rice (*OsCBFIV-1B.1*) and sugarcane (*SoCBFIV*) representatives were found to be distantly related to the core group 4 genes suggesting that the ancestral group 4 gene continued to evolve following the divergence of the corresponding plant subfamilies. A similar observation was noted previously by Skinner et al. (2005). Based on the present data, barley has at least 7 genes in group 4 (4 original plus 3 specifically amplified in barley) compared to 9 possible genes in wheat.

In conclusion, these analyses reveal that hexaploid wheat contains at least 15 *CBF* genes, and could contain up to 23 to 25 *CBF* genes. The latter estimates assume hexaploid wheat has retained orthologs of all *T. monococcum CBF* genes and of *HvCBFIa-1* and *OsCBFI-1F*. The *Poaceae CBF* genes can be divided into 10 groups with six of these (CBFIIIc, IIId, IVa, IVb, IVc and IVd) having evolved only in the *Pooideae*.

Bioinformatic analysis of wheat CBF proteins

The *TaCBF* genes identified in this study encode for proteins ranging from 202 to 290 amino acids that share homology with other CBF proteins. Analysis of these protein sequences reveals that they contain, to different degrees, the characteristic motifs found in this family (Wang et al. 2005; Skinner et al. 2005; Nakano et al. 2006). From the N- to C-terminus, we can identify the AP2 DNA-binding domain flanked by the CMIII-3 motif, and then the CMIII-1 motif, the CMIII-2 motif, and the conserved C-terminus LWSY motif (or CMIII-4). Analysis of the ORF of homeologous copies reveals that four groups have a member with a sequence difference at the C-terminus. The *Ta*CBFIVa-2.2 protein contains an extension of 30 amino acids past the last motif because of a T→C transition that

destroys a termination codon. In the *TaCBFIIIc-3.1* gene, a transversion in the last codon creates a premature termination codon and truncates the motif to LWS. The *Ta*CBFIVb-A20 protein contains an extension of 5 amino acids comparatively to other members of this group because of a deletion of the termination codon. In the *TaCBFIVd-B22* gene, a 38 base insertion downstream of the region encoding CMIII-1 motif changes the reading frame in the remainder of the sequence thus destroying the motifs CMIII-2 and -4. How these changes impact the activity of these proteins remains to be established.

Previously, Wang et al. (2005) had noted that the hydrophobic clusters (HCs) in the activation domain of CBF proteins were evolutionarily conserved and demonstrated their functional importance and redundant nature. To identify some of the structural differences associated with the 10 CBF groups classified by phylogenetic analysis, HC analysis was performed to highlight changes to internal faces of regular secondary structures (α-helices or β-strands) that may impact their functional activity. In groups that contain only one or two members, rice and barley orthologs were included in the analysis to examine the extent of structure conservation over evolutionary time. Two regions were analyzed and include the AP2 DNA-binding domain with the motifs CMIII-3 and CMIII-1 flanking it, and the C-terminal activation domain region containing motifs CMIII-2 and CMIII-4. For comparative purposes, we included *Rh*CBFI-1 and *Zo*CBFI-1 from the monocotyledonous orders *Arecales* and *Zingiberales*, respectively, the lone rice group I protein *Os*CBFI-1F, and *At*CBF3 from the eudicot *Arabidopsis thaliana*.

Analysis of the AP2 DNA-binding region reveals that the number/positions of HCs are relatively well conserved in the different CBF protein groups (Figure 2). These clusters correspond to previously defined regions involved in β-strand and α-helix formation in the AP2 domain (Allen et al. 1998). HC1 and HC5 show a high conservation in their length among the different groups. Some small qualitative differences are observed in the AP2 DNA-binding region when comparing the 10 groups of CBF proteins. For example, HC2 is on average larger in CBFIV groups, decreases in size in CBFIII groups, and is smallest in CBFIa and CBFII groups while

an inverse relationship is seen for HC4. The HC2 and HC3 regions were found to be the most useful in specifically defining the four CBFIV groups. The CBFIVa group specifically contains the largest HC3 while the CBFIVd group contains the largest hydrophobic character around the HC2 region. In the case of CBFIVb, HC2 is extended more towards HC1 comparatively to the CBFIVc group where it is extended towards HC3. The presence of glycine (G) and proline (P) residues is also a characteristic that can differentiate CBF groups. The P/G pattern between HC1 and HC2 is a characteristic of CBFIa proteins while the one between HC2 and HC3 is strictly conserved in the four CBFIII groups and the remaining group III rice proteins. The proline between HC2 and HC3 is lacking only in groups CBFIVb and CBFIVc.

Analysis of the regions surrounding the AP2 DNA-binding domain reveals significant differences between different CBF groups (Figure 2). In the CMIII-1 motif, the presence and length of HC and the P/G pattern are capable of differentiating between the CBFIa and CBFII groups, between the CBFIIIa and the remaining CBFIII groups, and between the four CBFIV groups. In the CMIII-3 region, the absence of a P is specific for the CBFIa and CBFIVa groups, and a larger HC defines groups CBFIa, CBFII and CBFIVa. These results show that all CBF groups besides CBFIIIb, CBFIIIc and CBFIIId may have slightly different folding of their secondary and/or tertiary structures of their DNA-binding domains.

Analysis of the C-terminal activation domain region (CMIII-2 and CMIII-4 motifs) reveals that CBFIa and CBFIII proteins contain fewer but larger HCs compared to CBFII and CBFIV proteins which contain a greater number of shorter HCs (Figure 3). In the vicinity of motif CMIII-2, CBFIa proteins contain only one HC bordered by P (Figure 3). This structure is also found in group CBFIIIa. However, this group can be differentiated from group Ia by a different G pattern in motif CMIII-2, the absence of a HC upstream of CMIII-2, and the only CBFIII group to contain a P in motif CMIII-4. CBFIIIb and CBFIIId groups contain a larger HC than CBFIIIa, and share a similar structural pattern from the CMIII-2 motif to the end of the protein. However, CBFIIId group can be differentiated by the presence of a HC upstream of CMIII-2. Unique features that characterize the CBFIIIc group include a P

that creates two HCs within motif CMIII-2, and an additional HC downstream of this motif.

The CBFII and CBFIV groups contain between 4 to 6 HCs in their C-terminal activation region which is similar to the number identified in *Arabidopsis* (Wang et al. 2005). However, the structural patterns are different. In fact, the lone protein *Os*CBFI-1F is the one that shows the highest structural similarity with *At*CBF3 suggesting that it may represent a closer version of the ancestral type CBF in monocots. The CBFII group contains 5 HCs and is unique among the groups analyzed since it does not contain P at the end of motif CMIII-2. The CBFIVa and CBFIVd groups share a similar structure with 3 HCs in motif CMIII-2 and up to 5 common HCs in total. Group CBFIVd can be differentiated by the presence of the first HC upstream of motif CMIII-2 in all members and a P in motif CMIII-4. Groups CBFIVb and CBFIVd contain only two HCs in motif CMIII-2, and these display variable lengths. In addition, group CBFIVc contains an additional HC downstream of CMIII-2 which is not found in any other group IV CBF. In summary, these results show that most groups described in this study display some structural differences in the AP2 DNA-binding region while all groups show differences in their C-terminal activation regions. An important observation from these results is that rice proteins (*Os*CBFIII-1D, *Os*CBFIII-1I, *Os*CBFIII-1J and *Os*CBFIV-1B.1) do not show clear structure conservation with any of the described groups, corroborating their non-orthologous relationship. This is in contrast to the rice members of groups CBFIa, CBFII and CBFIIIa and partly *Zo*CBFI-1 and *Rh*CBFI-1 that do show structures that are orthologous in nature, and have been conserved since the divergence of the respective branches.

Expression of wheat CBF genes

CBF genes are known to be induced rapidly upon exposure to LT. To determine the expression behaviour of the 15 *TaCBF* gene groups identified in this study, we initiated a LT time course using two cultivars differing in their FT capacities. The LT treatment was initiated 4 hours after dawn and continued for 7

days. Probes outside of the AP2 DNA-binding domain were designed for Northern blot analyses to avoid cross reaction with known *TaCBF* genes (Table S3). The results obtained with *TaCBFIVb-21.1* and the genes from groups *CBFIa* and *CBFIIIc* are not included in Figure 4 since no signals were detected. This suggests that these genes are expressed at low levels under the conditions assayed. The remaining 11 genes displayed signals and are shown in Figure 4. Analysis of these results allows two general observations to be drawn. The first is that all assessed and detectable *TaCBF* genes display a transient induction profile. In fact, all *TaCBF* genes showed little or no expression before the onset of the LT treatment. They were induced by LT and attained maximum levels after 4 to 6 hours of treatment, and then returned to basal levels after 1 to 7 days of treatment. The second observation revealed by our Northern analyses is that 9 of the 11 *TaCBF* genes assayed are expressed to higher levels in the winter cultivar compared to the spring cultivar suggesting that higher *TaCBF* expression is associated with the winter cultivar's superior FT development capacity. However, there are differences in the quantitative accumulation of certain *TaCBF*s. For example, the expression of *TaCBFIIId-B12* and *TaCBFIIId-15.2* was not detected in the spring cultivar while the remaining 7 genes were expressed at low levels compared to the winter cultivar.

Since the northern expression study on *TaCBF*s (Figure 4) was not sensitive enough to measure the basal expression of *TaCBF*s, we decided to use the more sensitive quantitative real-time PCR to quantify the initial LT response in a winter and spring cultivar. The experiment was designed to include several consecutive 20 and 4°C time points for evaluating the extent of basal versus LT inducible fluctuation, and to initiate the LT treatment near the end of the day to better reflect natural conditions. This experimental design also allowed a comparison of the influence of day period on *TaCBF* induction by LT. Primer sets used in these experiments were designed against only one copy of each of the 15 *TaCBF* gene groups identified (Table S4) and shown to be specific by mapping these genes to unique chromosome arms (Table 1). To compare the *CBF* induction patterns in a winter and spring cultivar, panels in Figure 5 were generated using the independent calibrators winter

08:00 and spring 08:00, respectively. This resulted in y-axis scales that are not directly comparable among cultivars but allow the direct visualization of the LT effect on gene expression. Once these profiles were obtained, the relative quantity of winter versus spring LT accumulation was experimentally determined after two hours of LT exposure at 22:00 (shown as the winter/spring expression at 22:00 in Figure 5) while the basal expression value was measured for the control point 18:00 (results not shown). From the expression pattern of *TaCBF* genes, 13 showed statistically significant basal and/or LT expression (Figure 5). The exceptions that did not show any reproducible signal were *TaCBFIa-A11* and *TaCBFIIIc-B10* in both cultivars (results not shown).

Results of Figure 5 show that in almost all cases the maximum accumulation of *TaCBF*s occurs after two hours of LT treatment. These results are in contrast with those of 4 and 6 hours observed in Figure 4 and suggest that the pattern of induction is influenced by the period of the day when the LT treatment was initiated. Several other observations can be noted from the results presented in Figure 5. *TaCBFIIIc-D3* displays an extreme transient expression profile showing low expression in all points examined except the 2 hour LT treatment. A similar expression profile was observed for the barley ortholog (Choi et al. 2002) suggesting evolutionary conservation of regulation pattern. The fact that no detectable expression of *TaCBFIa-A11* and *TaCBFIIIc-B10* was observed in this study suggests that *CBFIa* and *CBFIIIc* genes may be expressed under specific conditions and/or at extremely low levels. Other observations include: a near constitutive expression for *TaCBFII-5.2*, a low and transient induction for *CBFIIIa* and *CBFIVb* genes, and a high and more sustained expression for most *CBFIIId*, *CBFIVc* and *CBFIVd* genes. A comparison of winter and spring LT induction profiles reveals that they are qualitatively very similar. The quantitative comparison of the 2 hour LT time points reveals that *CBFII*, *CBFIIIa* and *CBFIIIc* genes are expressed to similar levels (within a three fold factor) in both cultivars. On the other hand, *CBFIIId*, *CBFIVa*, *CBFIVb*, *CBFIVc* and *CBFIVd* genes (except *TaCBFIVd-D9*) show increased LT expression (4.7 fold and more) in the winter compared to the spring cultivar. In addition, these 5 groups (except for

TaCBFIIId-B12) also show a higher basal expression in the winter cultivar (results not shown), and this can be easily evaluated for some members when comparing the induction profiles in Figure 5 if we consider that the LT expression of these genes is more than 4.7 fold higher in winter compared to spring cultivars as indicated above. These results indicate that the 5 CBF groups are associated with both the superior inherited and LT inducible capacities of the winter cultivar to develop FT.

An additional interesting observation that emerged from this experiment is that the expression of several *TaCBF* genes was not constant during the seven 20°C time points indicating that cold treatment is not needed for their expression. The expression of these genes was high in the vicinity of 8 to 10 hours after dawn and then decreased in both winter and spring cultivars (for example see *TaCBFIVb-D20* in Figure 5). To better visualize this behaviour, the 20°C experiment was repeated with the cultivar Norstar for a full 24 hour period without cold treatment (Figure 6). These results reveal that genes from the four CBFIV groups and 2 of the 3 genes from the CBFIIId group show a diurnal fluctuation in their expression with maxima appearing between 8 and 14 hours after dawn and minima between 20 and 24 hours after dawn under long day conditions. This diurnal fluctuation is reproducible since it was observed for two additional cycles in experiments with *TaCBFIVb-D20*, *TaCBFIVc-B14*, *TaCBFIVd-B4*, *TaCBFIVd-D9* and *TaCBFIVd-B22* (results not shown). Therefore, the 5 groups that were found to be more expressed in the winter cultivar also display a characteristic diurnal fluctuation during growth at 20°C.

Discussion

Wheat is a good model species for studying FT since its tolerance lies between that of freezing sensitive plants (rice, maize and oat) and the extremely tolerant species rye. To decipher the genetic basis underlying the different capacities of temperate cereal species in developing FT, we have initiated the identification of genes that have the potential to influence FT. Since *CBF* genes have been widely implicated in cold acclimation in many species, we identified and characterized 15 *TaCBF* gene groups in hexaploid wheat. Our analyses revealed that wheat species, *T. aestivum* and *T. monococcum*, have a large and complex *CBF* family with up to 25 different *CBF* genes. The large number of *CBF* genes in wheat is comparable to the number found in barley (20 or more) (Skinner et al. 2005) but contrasts with the 10 genes present in rice and 6 in *Arabidopsis*. It is not known why freezing tolerant cereals have evolved and maintained so many *CBF* genes. Since the amplification of the *CBF* gene family has evolved independently after the monocot-eudicot divergence (Qin et al. 2004; Bräutigam et al. 2005; Xiong and Fei 2006), a thorough characterization of a large number of *CBF* genes in wheat seemed a daunting task. At present, it is difficult to assess if all *TaCBF*s have specific functions, obtained by subfunctionalization and neofunctionalization, or if they all have redundant functions. Therefore, an objective of this study was to classify *CBF* genes into structuralcategories that would help orient functional studies. Towards this goal, we studied the evolution of the *CBF* family in monocotyledons and determined their structural characteristics using HCA to display conservation/changes that could affect protein secondary structure.

This study indicates that the *CBF* amplification seen in wheat has occurred quite recently and followed the emergence of the *Oryzaceae*, *Panicoideae* and *Pooideae* lineages. From the phylogenetic analysis, it is easily observable that these subfamilies had already evolved representatives of groups Ia, II, IIIa and possibly IIIb suggesting that orthologs within these groups should have common functions. The

HCA analyses confirmed that rice and wheat orthologs have similar structural characteristics that have been conserved over evolution. On the other hand, groups IIIc, IIId, IVa, IVb, IVc and Ivd, representing 18 genes in wheat species, arose following the emergence of the *Pooideae* since no rice or maize *CBF*s clustered within these groups. After the emergence of *Oryzaceae*, rice only gained 4 genes. However, these rice genes are distantly related to the groups containing wheat genes and their proteins do not display similar structural features. Therefore, the 18 wheat genes in groups IIIc, IIId, IVa, IVb, IVc and Ivd, and the 4 unclassified group III and IV rice genes may represent the *CBF* response machinery that evolved in *Pooideae* and *Oryzaceae*, respectively, as they radiated into specific habitats. Since these arose after the subfamilies split, it is not surprising to see that structural patterns are not conserved in the respective CBF proteins. There is some indications that the *CBF* family is still (or has been recently) evolving under some selective pressure. The first comes from the observation that, although the evolutionary distance between groups IVa, IVb, IVc and IVd is small, their proteins display notable structural differences in their AP2 domain and C-terminal region. This is evident between groups but can also be seen between members of group IVa (Figure 2) and IVb (Figure 3). Finally, the observation that barley contains several duplicated genes with high similarities (Skinner et al. 2005) suggests that they arose after the wheat-barley divergence with some (those with identities above 95%) having arisen in the last 4 MY. As more *CBF* sequence information becomes available, it will allow a detailed evaluation of the number and nature of *CBF* gene groups found in different subfamilies of the *Poaceae* and even in other monocot orders. In a general sense, this will lead to a better understanding of the evolution of CBF-mediated tolerance to abiotic stresses, and in a more practical way, it may allow associating the emergence (or loss) of certain *CBF* genes (or groups) with maximum species freezing tolerance.

An interesting conclusion that can be drawn from the phylogenetic study is that the evolution of the *Pooideae* in a specific ecosystem has impacted the CBF signaling machinery by increasing the total number of *CBF* genes and the number of structural categories, as also corroborated by the HCA analyses. This complexity is

found in the *Triticeae* tribe and suggests that these plants have faced a strong selective pressure to maintain genes that will help them to perform well under a variety of environmental conditions. The selection pressure may not be constant but could intensify during generations that experience an unusually severe winter.

Studies of monocot CBFs have shown that they share the conserved domains present in this protein family, and that they are capable of binding a DRE-related *cis* element and inducing *COR* gene expression in *Arabidopsis* although with a lesser efficiency or an incomplete response (Dubouzet et al. 2003; Qin et al. 2004; Skinner et al. 2005) compared to overexpression of endogenous *Arabidopsis CBF*s (Jaglo-Ottosen et al. 1998; Liu et al. 1998). This can be partly explained by structural differences which make monocot CBFs less efficient in replacing the endogenous *Arabidopsis* protein function. The independent evolution of *CBF* genes in plants will certainly make it harder to elucidate the exact roles of cereal genes in model species like *Arabidopsis* that are more evolutionary distant. Therefore, to understand the exact functions of *Pooideae*- and possibly *Triticeae*-specific groups/genes, it will be essential to study these *CBF* genes in species such as wheat and barley. In addition, determining the exact contributions of members of the *CBF* family may be even more complex than anticipated since members of other subgroups of group III ERF proteins such as *Hv*CBF7 from barley (Skinner et al. 2005) and TINY2 from *Arabidopsis* (Wei et al. 2005) have been shown to be cold-regulated and capable of binding a DRE cis element. Their contribution to the regulation of *COR* gene expression in species with large *CBF* families or their possible compensation in species with small families needs to be explored.

The HCA analysis of the AP2 DNA-binding domain was capable of differentiating 7 out 10 groups suggesting that protein structure and binding properties may be affected. Results of Xue (2002) demonstrated that *Hv*CBFIa-1 preferred TTGCCGACAT as a binding site while *Hv*CBFIVa-2 (Xue 2003) preferred YYGTCGACAT. In addition, *Hv*CBFIVa-2 and *Hv*CBFIVd-4A showed a LT dependence for maximal binding activity (Xue 2003; Skinner et al. 2005). These results corroborate that functional differences can be visualized through HCA

analyses and allow us to predict that up to 7 groups may show some differences in DNA-binding properties and that group members will show a redundancy in binding. Determining such properties will be important for understanding the possible differences/overlap in regulons controlled by CBF groups. It was recently suggested that small variation in transcription factor binding consensus could have important consequences for bioactivity (Benedict et al. 2006) and some of the barley CBFs have been demonstrated to have differential affinity for specific CRT/DRE motifs (Xue 2002; Xue 2003; Skinner et al. 2005). On the other hand, the HCA analysis of the C-terminal activation domain revealed substantial differences between the 10 groups. The structural patterns were relatively well conserved between group members even for those corresponding to distantly related species. The varied patterns detected may be at the base of structural differences which could impact protein folding, recognition of specific interaction partners and the transactivation potential of a specific set of CBF proteins. The molecular dissection of *Arabidopsis* CBF1 (Wang et al. 2005) showed that certain hydrophobic clusters and other structural determinants in this region were important in regulating transactivation potential of this region. In *Brassica napus*, two closely related CBF proteins were shown to have substantially different transactivation potentials in yeast and tobacco (Zhao et al. 2006). The major differences between these proteins lie in the C-terminal activation domain within the hydrophobic clusters. Such behaviour has not been reported for the *CBF1*, *2*, and *3* genes of *Arabidopsis* (Gilmour et al. 2004) which diverged from *Brassica* some 24 MYA (Koch et al. 2000) suggesting that even related species may have evolved some specific differences in CBF properties. These examples illustrate that the C-terminal activation domain of different groups could have distinct properties that play specific roles. In addition, this last observation suggests that some CBF properties in group IVa (Figure 2) and IVb (Figure 3) may even differ between the closely related species barley and wheat. The three structurally different groups IIIc, IIId and IVd share the characteristic of having several members with similar structural patterns which contrast with the lower complexity present in other groups. An explanation for the selection and conservation of duplicated genes that seem to

accomplish redundant functions comes from work in yeast and suggests a selection for a higher flux through the pathways controlled by these genes (Papp et al. 2004). Therefore, the wheat genes from these groups may be necessary in certain circumstances where maximal induction is needed to achieve very high levels of LT tolerance and/or activation of a large number of genes in a regulon.

Quantitative RT-PCR analyses revealed that the expression level of groups IIId, IVa, IVb, IVc and IVd was more pronounced in the winter cultivar (4-fold and more) compared to the remaining groups assayed (3-fold and less). This was also demonstrated in barley and wheat for members of the *CBFIV* groups (Kume et al., 2005; Skinner et al. 2005). It is probably not a coincidence that 5 of the 6 groups that evolved in the *Pooideae* show expression levels that are correlated with the winter cultivar's capacity to develop LT tolerance. Previous studies had already noted that the majority of the expansion of *CBF* genes has occurred from an ancestral cluster/locus (Skinner et al. 2005; Miller et al. 2006; Skinner et al. 2006). In rice, three *CBF* genes (*OsCBFIIIa-1A*, *OsCBFIIIb-1H* and *OsCBFIV-1B.1*) are present as a tandem cluster on a region on rice chromosome 9 that is collinear with the chromosome 5 region of the *Triticeae* where the *CBF*s occur. Therefore in the *Triticeae* (*T. monococcum*, barley and hexaploid wheat), there has been amplification of the genes in this region to give rise to the 6 groups present specifically in this tribe. These observations suggest that as FT-associated *CBF* genes were amplified, selective pressure has maintained their role in FT. The genetic capacity of the winter cultivar to induce a higher level of expression during the LT response is also reflected in the constitutive levels measured under control growth conditions. Since LT is not involved in this regulation, the higher level in winter versus spring wheat comes from different inheritable capacities to express *CBF* genes. A similar association was observed between the inheritable level of *CBF* expression and the LT tolerance of different *Arabidopsis* lines collected at different latitudes (Hannah et al. 2006). These observations are not surprising since a study had already shown an association between higher levels of the WCS120 protein family and cultivar capacity to develop FT (Houde et al. 1992).

The expression patterns also revealed that groups IIId, IVa, IVb, IVc and IVd displayed a diurnal fluctuation that peaked 8 to 14 hours after dawn. This natural rhythm influenced the time course of LT induction with higher inductions during evenings (maximum at 2 hours) and lower induction during mornings (maximum at 4 to 6 hours). Although the peak period of induction does not coincide with the coolest period of the day, it does coincide with the daily decrease of temperature during sunset suggesting the rhythm is preceding/anticipating the event. Circadian clock regulation of *CBF* expression has been described in *Arabidopsis* (Fowler et al. 2005), and therefore, may be common in plants capable of developing FT since it would confer a selective advantage during sudden drops in LT. In temperate cereals, group IV proteins (*Hv*CBFIVa-2 and *Hv*CBFIVd-4A) were shown to bind *cis* elements in a LT-dependent manner. If this property is present in all group IV members of temperate cereals, the daily accumulation of these proteins during normal growth conditions would not cause profound changes in *COR* gene expression and thus prevent wasting cellular resources since their DNA-binding activity is relatively low at this temperature. Once exposed to a sudden drop in LT, the DNA-binding activity of these factors would increase, and this would immediately impact *COR* gene expression. The accumulation of partially inactive factors during warm growth conditions would also alleviate deleterious symptoms from developing as observed in transgenic plants constitutively overexpressing complete or portions of *CBF* genes (Liu et al. 1998; Wang et al. 2005; Ito et al. 2006). Therefore, group IV proteins from temperate cereals may ultimately represent a uniquely engineered protein group that functions as a first line of protect against sudden drops in LT. The functional studies of group IV CBFs is thus essential to understand the specific contribution of these groups and their impact on the range of LT tolerance capacities observed in temperate cereals. In addition, these studies may identify unique properties that could be incorporated in CBF proteins from other species.

In conclusion, this study revealed that wheat species, *T. aestivum* and *T. monococcum*, may contain up to 25 *CBF*s. These genes can be divided into at least 10 groups that share a common phylogenetic origin and similar structural characteristics.

Six of these groups (CBFIIIc, IIId, IVa, IVb, IVc and IVd) are found only in the *Pooideae*, suggesting that they evolved recently during the colonization of temperate habitats. Expression studies revealed that 5 groups (CBFIIId, IVa, IVb, IVc and IVd) display higher constitutive and LT-inducible expressions in the winter cultivar. The higher inherited and inducible *CBF* expression suggests that these groups may be major components that regulate the capacities of *Pooideae* species to develop LT tolerance.

Acknowledgement We wish to thank F. Allard, M. Champoux, N. Toudji and K. Tremblay for their technical assistance. The excellent editing work of one of the anonymous reviewers is also appreciated. This study was supported by Genome Canada and Génome Québec grants (F.S. and M.H.), and by Natural Sciences and Engineering Research Council of Canada discovery grant (F.S. and J.D.). M.B. was supported by a scholarship from the Ministry of State for Higher Education and Scientific Research, Egypt.

Tables

II-Table 1: Nomenclature and characteristics of wheat *CBF* genes isolated from the cultivar Norstar

Gene name	Accession number	Isolation method[a]	Chromosome location[b]	cDNA (bp)[c]	Open reading frame (CDS)	Total amino acids
TaCBFIn-A11	EF028751	3	2AL	962	54–710	218
TaCBFII-5.1	EF028752	3		1,069	98–775	225
TaCBFII-5.2 [d]	EF028753	3		991	109–768	219
TaCBFII-5.3	EF028754	3		876	107–793	228
TaCBFIIIa-6.1	EF028755	1		958	85–795	236
TaCBFIIIa-6.2	EF028756	3		947	112–840	242
TaCBFIIIa-D6	EF028757	2	5DL-1; 5DL-2	959	107–823	238
TaCBFIIIc-3.1	EF028758	3		896	<1–708	>235
TaCBFIIIc-3.2	EF028759	3		922	111–851	246
TaCBFIIIc-D3	EF028760	1	5DL-1; 5DL-2	963	105–842	245
TaCBFIIIc-B10	EF028761	1, 3	5B	1,000	91–813	240
TaCBFIIId-12.1	EF028762	3		1,033	130–867	245
TaCBFIIId-B12	EF028763	1, 3	5B	970	88–825	245
TaCBFIIId-A15	EF028764	3	5AL-10; 5AL-8	1,010	107–826	239
TaCBFIIId-15.2	EF028765	3		926	110–835	241
TaCBFIIId-A19	EF028766	2, 3	5AL-10; 5AL-8	984	184–888	234
TaCBFIIId-B19	EF028767	1	5B	954	150–854	234
TaCBFIIId-D19	EF028768	1	5DL-1; 5DL-2	965	166–870	234
TaCBFIVa-A2	EF028769	2, 3	5AL-10; 5AL-8	950	52–729	225
TaCBFIVa-2.2[e]	EF028770	3		822	<3–693	>230
TaCBFIVa-2.3	EF028771	3		660	86–>660	>191
TaCBFIVb-A20	EF028772	3	5AL-10; 5AL-8	1,077	77–730	217
TaCBFIVb-B20	EF028773	1	5B	887	63–701	212
TaCBFIVb-D20	EF028774	2, 3	5DL-1; 5DL-2	980	77–715	212
TaCBFIVb-21.1	EF028775	3		979	193–801	202
TaCBFIVb-D21	EF028776	3	5DL-1; 5DL-2	1,066	380–988	202
TaCBFIVc-14.1	EF028777	1, 3		1,003	120–758	212
TaCBFIVc-B14	EF028778	2	5B	863	106–750	214
TaCBFIVc-14.3	EF028779	1		892	119–763	214
TaCBFIVd-4.1	EF028780	1		872	53–721	222
TaCBFIVd-B4	EF028781	1	5B	865	56–724	222
TaCBFIVd-9.1	EF028782	2		1,080	99–908	269
TaCBFIVd-B9	EF028783	2	5B	999	77–886	269
TaCBFIVd-D9	EF028784	2, 3	5DL-1; 5DL-2	1,063	110–919	269
TaCBFIVd-A22	EF028785	3	5AL-10; 5AL-8	1,219	95–922	275
TaCBFIVd-B22[f]	EF028786	1, 3	5B	1,252	94–966	290
TaCBFIVd-D22	EF028787	3	5DL-1; 5DL-2	1,211	96–923	275

[a] The methods used to isolate *CBF* genes in wheat are described by the codes: *1* WGAS EST; *2* cDNA library screening; *3* PCR amplification from cDNA libraries

[b] Indicates either a location to a chromosome, a chromosome arm or a deletion breakpoint interval

[c] The total number of bases in a cDNA does not include the poly A tail and gene specific primer sequences if they were not confirmed by additional sequence information

[d] Mapping did not produce a chromosome location. However, orthologs in *T. monococcum* (Miller et al. 2006) and barley (Skinner et al. 2006) were mapped to chromosomes 7A and short arms of 7H, respectively

[e] Contains a T→C transition that destroys a termination codon and extends the coding sequence for 93 bases

[f] Contains a 38 base insertion that creates a frame shift for the remaining coding sequence

II-Table 2: List of monocotyledon CBF genes and their proposed nomenclature

Plant species[a]	Proposed gene name	Gene name	Accession number	Type of sequence
Poales/Poaceae/Oryzaceae				
Oryza sativa	*OsCBFI-1F*	*OsDREB1F*	AY785897	mRNA
	OsCBFIa-1G	*OsDREB1G*	AK060550	mRNA
	OsCBFIa-1E	*OsDREB1E*	AY785896	mRNA
	OsCBFII-1C	*OsDREB1C*	AY327040	mRNA
	OsCBFII-1D	*OsDREB1D*	AY785895	mRNA
	OsCBFIII-1I	*OsDREB1I*	XM_483622	Genomic
	OsCBFIII-1J	*OsDREB1J*	XM_483621	Genomic
	OsCBFIIIa-1A	*OsDREB1A*	AF300970	mRNA
	OsCBFIIIb-1H	*OsDREB1H*	AP008215 (CDS: 20099705-20098965)	Genomic
	OsCBFIV-1B.1	*OsDREB1B.1*	AY785894	mRNA
Poales/Poaceae/Pooideae				
Agrostis capillaris	*AcCBFIIIa-6*		DV858477	EST
	AcCBFIIId-19		DV853050	EST[b]
Agrostis stolonifera	*AstCBFIIIc-3*		DY543542	EST[b]
Avena sativa	*AsCBFIIId-12*	*CBF4*	AM071409	mRNA
	AsCBFIIId-16A	*CBF1*	AM071406	mRNA
	AsCBFIIId-16B	*CBF2*	AM071407	mRNA
	AsCBFIVa		AM071408	EST[b]
Brachypodium distachyon	*BdCBFIa-1*		DV479443	EST
	BdCBFII-5		DV482761, DV481106, DV485162	EST
	BdCBFIIIa-6		DV485858, DV489297, DV487222, DV478843, DV489355, DV484024, DV478031	EST
Festuca arundinacea	*FaCBFIIIa-6*	*DREB1A*	AJ717399	mRNA
	FaCBFIIIc-3		DT710563, DT701428	EST
	FaCBFIVa-2	*DREB1*	AY423713	mRNA
Hordeum brevisubulatum	*HbCBFIVa-2*	*DREB1*	DQ250027	mRNA
Hordeum vulgare	*HvCBFIa-1*	*HvCBF1*	AY785836	Genomic
	HvCBFIa-11	*HvCBF11*	AY785890	Genomic
	HvCBFII-5	*HvCBF5*	AY785855	Genomic
	HvCBFIIIa-6	*HvCBF6*	AY785860	Genomic
	HvCBFIIIc-3	*HvCBF3*	AY785845	Genomic
	HvCBFIIIc-8A	*HvCBF8A*	AY785868	Genomic
	HvCBFIIIc-10A	*HvCBF10A*	AY785882	Genomic
	HvCBFIIIc-13	*HvCBF13*	DQ095158	Genomic
	HvCBFIIId-12	*HvCBF12*	DQ095157	Genomic
	HvCBFIVa-2A	*HvCBF2A*	AY785841	Genomic
	HvCBFIVc-14	*HvCBF14*	DQ095159	Genomic
	HvCBFIVd-4A	*HvCBF4A*	AY785849	mRNA
	HvCBFIVd-9	*HvCBF9*	AY785878	Genomic
Lolium perenne	*LpCBFIIIa-6*	*CBF3*	AY960831	mRNA
Secale cereale	*ScCBFIVb-20*	*Clone 79 CBF*	AF370728	mRNA
	ScCBFIVd-9A	*Clone 80 CBF*	AF370729	mRNA
	ScCBFIVd-9B	*Clone 81 CBF*	AF370730	mRNA
Triticum aestivum	*TaCBFIa-A11*		EF028751	mRNA
	TaCBFII-5.1		EF028752	mRNA
	TaCBFIIIa-6.1		EF028755	mRNA
	TaCBFIIIc-3.1		EF028758	mRNA
	TaCBFIIIc-B10		EF028761	mRNA
	TaCBFIIId-12.1		EF028762	mRNA
	TaCBFIIId-A15		EF028764	mRNA
	TaCBFIIId-A19		EF028766	mRNA
	TaCBFIVa-A2		EF028769	mRNA
	TaCBFIVb-A20		EF028772	mRNA
	TaCBFIVb-21.1		EF028775	mRNA
	TaCBFIVc-14.1		EF028777	mRNA
	TaCBFIVd-4.1		EF028780	mRNA
	TaCBFIVd-9.1		EF028782	mRNA
	TaCBFIVd-A22		EF028785	mRNA

II-Table 2: continued

Plant species[a]	Proposed gene name	Gene name	Accession number	Type of sequence
Triticum Monococcum	*TmCBFII-5*	*TmCBF5*	AY951947	Genomic
	TmCBFIIIb-18	*TmCBF18*	AY951946	Genomic
	TmCBFIIIc-10	*TmCBF10*	AY951950	Genomic
	TmCBFIIIc-13	*TmCBF13*	AY951951	Genomic
	TmCBFIIId-16	*TmCBF16*	AY951944 (CDS: 181665-182366)	Genomic
	TmCBFIIId-17	*TmCBF17*	AY951945 (CDS: 91223-90357)	Genomic
	TmCBFIVa-2	*TmCBF2*	AY951945 (CDS: 22116-22838)	Genomic
	TmCBFIVd-4	*TmCBF4*	AY951945 (CDS: 35360-34722)	Genomic
Poales/Poaceae/Panicoideae				
Panicum virgatum	*PvCBFIa-11*		DN144490, DN145877	EST
	PvCBFII-5		DN143696, DN145297, DN143313	EST
	PvCBFIIIa-6		DN144355, DN143145, DN143526	EST
Saccharum officinarum	*SoCBFIa-11*		CA212714, CA156028, CA080013, BQ534420	EST
	SoCBFII-5		CA162524, CA161788, CA089833, BQ533805	EST
	SoCBFIIIa-6		BU103690, CA090085, CA109731, CA105526	EST
	SoCBFIIIb		CA155733	EST[b]
	SoCBFIV		CA274081	EST[b]
Sorghum bicolor	*SbCBFII-5*	*SbCBF5*	AY785898	mRNA
	SbCBFIIIa-6	*SbCBF6*	AY785899	mRNA
Zea mays	*ZmCBFII-5*		DR964417, DV523865, DR790548, DV530932	EST
	ZmCBFIIIb-1A	*DREB1A*	AF450481	mRNA
	ZmCBFIIIb-1B		EB674710, DV539998, DV510250, DR821363, EB674709, EB400670, DR795602	EST
Arecales/Arecaceae/Coryphoideae				
Sabal palmetto	*SpCBFI-1*	*CBF-like*	DQ497730	Genomic
Rhapidophyllum hystrix	*RhCBFI-1*	*CBF-like 1*	DQ497742	mRNA
	RhCBFI-2	*CBF-like 2*	DQ497743	mRNA
Arecales/Arecaceae/Arecoideae				
Dypsis lutescens	*DlCBFI-1*	*CBF-like*	DQ497738	Genomic
Zingiberales/Zingiberaceae				
Zingiber officinale	*ZoCBFI-1*		DY354914, DY380878	EST
	ZoCBFI-2		DY344903, DY345420, DY380161, DY381223	EST

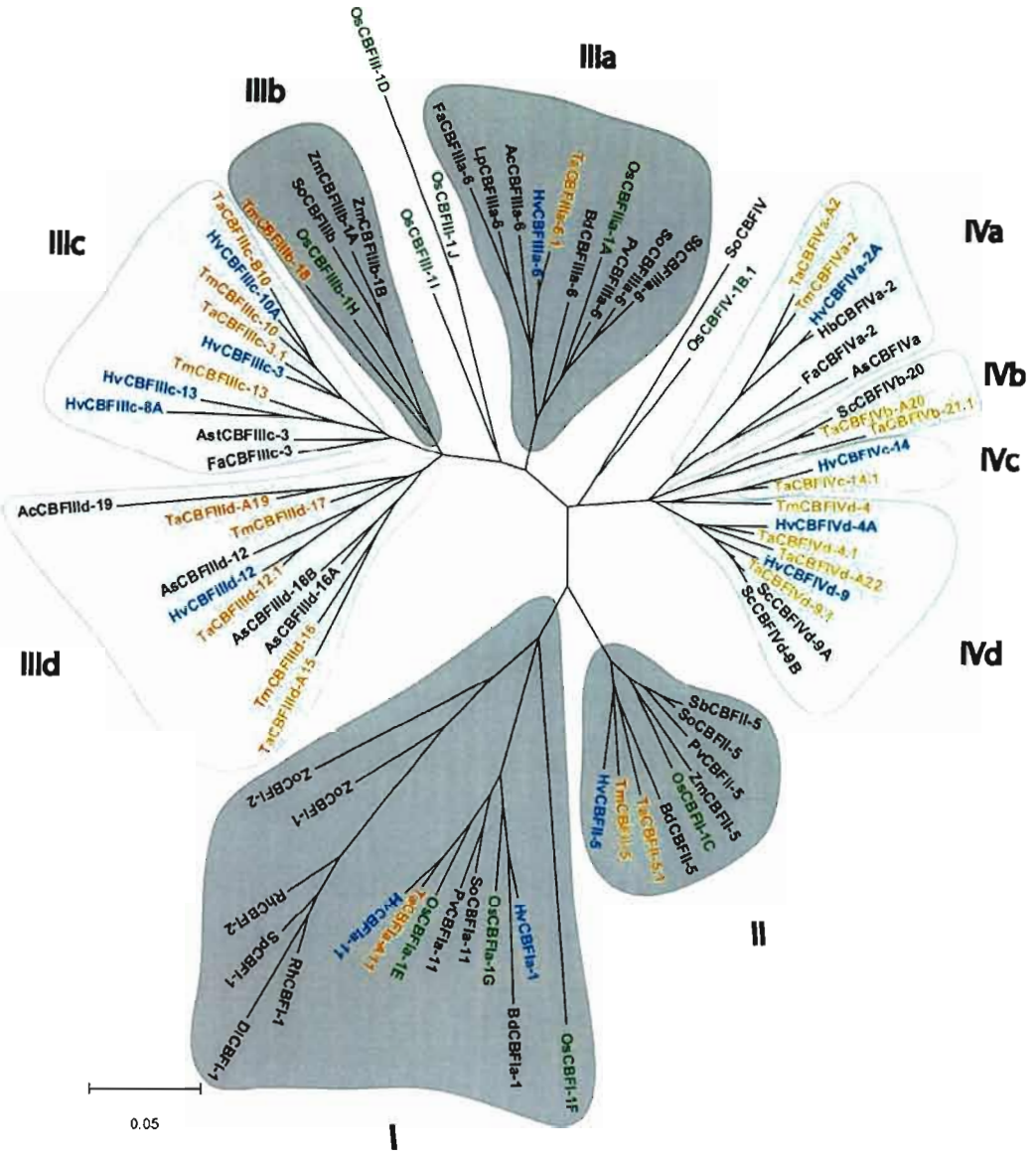
[a] Plant species are organized under the appropriate monocot order/family/subfamily

[b] The proposed gene names are tentative since they are based on partial EST sequences

Figures

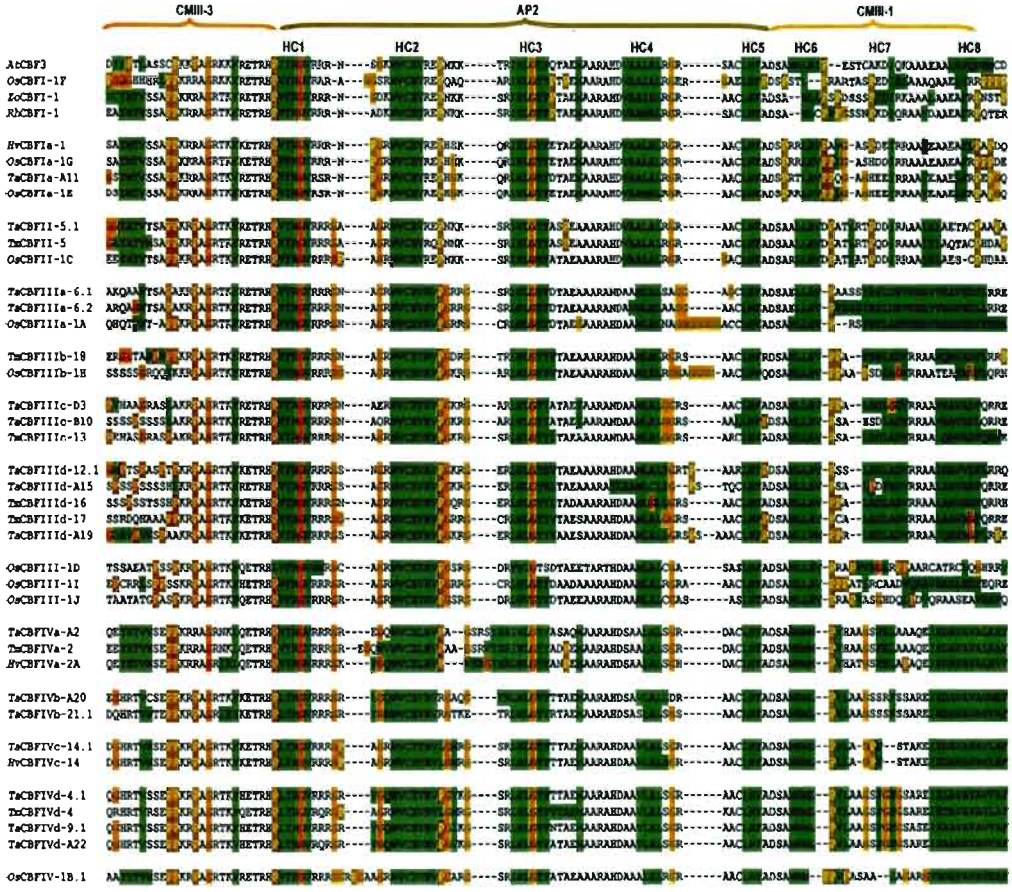
II-Figure 1: Phylogenetic relationships between monocot *CBF* genes.

The nucleotide sequence corresponding to the AP2 DNA-binding domain and the conserved flanking signature sequences PKK/RPAGRxKFxETRHP and DSAWR defined by Jaglo et al. (2001) were aligned using ClustalW and manually adjusted. An unrooted Minimum Evolution tree was derived from this alignment using the Kimura 2-parameter. The CBF nomenclature used is described in Table 2. *CBF* genes belonging to specific monophyletic groups are contoured. Complete CBF names are shown in red for wheat species, blue for barley, green for rice and black for all other species examined. Groups lightly shaded contain only *Pooideae* sequences. *TmCBFIVd-4* was included in group *IVd* based on the analysis of the complete sequence. Ac, *Agrostis capillaries*; As, *Avena sativa*; Ast, *Agrostis stolonifera*; Bd, *Brachypodium distachyon*; Dl, *Dypsis lutescens*; Fa, *Festuca arundinacea*; Hb, *Hordeum brevisubulatum*; Hv, *Hordeum vulgare*; Lp, *Lolium perenne*; Os, *Oryza sativa*; Pv, *Panicum virgatum*; Rh, *Rhapidophyllum hystrix*; Sb, *Sorghum bicolor*; Sc, *Secale cereale*; So, *Saccharum officinarum*; Sp, *Sabal palmetto*; Ta, *Triticum aestivum*; Tm, *Triticum monococcum*; Zm, *Zea mays*; Zo, *Zingiber officinale*


IIIa
IIIb
IIIc
IIId
IVa
IVb
IVc
IVd
I
II
OsCBFIII-1D
FaCBFIIIa-6
LpCBFIIIa-6
AcCBFIIIa-6
HvCBFIIIa-6.1
BdCBFIIIa-6
PvCBFIIIa-6
SbCBFIIIa-6
SoCBFIV
OsCBFIV-1B.1
TaCBFIVa-A2
TmCBFIVa-2
HvCBFIVa-2A
HbCBFIVa-2
FaCBFIVa-2
AsCBFIVa
ScCBFIVb-20
TaCBFIVb-A20
TaCBFIVb-21.1
HvCBFIVc-14
TaCBFIVc-14.1
TmCBFIVd-4
HvCBFIVd-4A
TaCBFIVd-4.1
TaCBFIVd-A22
HvCBFIVd-9
TaCBFIVd-9.1
ScCBFIVd-9A
ScCBFIVd-9B
SbCBFII-5
SoCBFII-5
PvCBFII-5
ZmCBFII-5
OsCBFII-1C
BdCBFII-5
TaCBFII-5.1
TmCBFII-5
HvCBFII-5
OsCBFII-1J
OsCBFII-1I
ZmCBFIIIb-1A
SoCBFIIIb
ZmCBFIIIb-1B
OsCBFIIIb-1H
TaCBFIIIc-B10
TmCBFIIIc-10
HvCBFIIIc-10A
TaCBFIIIc-10
TaCBFIIIc-3.1
HvCBFIIIc-3
TmCBFIIIc-13
HvCBFIIIc-13
HvCBFIIIc-8A
AstCBFIIIc-3
FaCBFIIIc-3
AcCBFIIId-19
TaCBFIIId-A19
TmCBFIIId-17
AsCBFIIId-12
HvCBFIIId-12
TaCBFIIId-12.1
AsCBFIIId-16B
AsCBFIIId-16A
TmCBFIIId-16
TaCBFIIId-A15
ZoCBFI-2
ZoCBFI-1
RhCBFI-2
SpCBFI-1
RhCBFI-1
DlCBFI-1
PhCBFIa-1.1
TaCBFIa-A1
SlCBFIa-1E
OsCBFIa-11
PvCBFIa-11
SoCBFIa-1G
OsCBFIa-1G
HvCBFIa-1
BdCBFIa-1
OsCBFI-1F
0.05

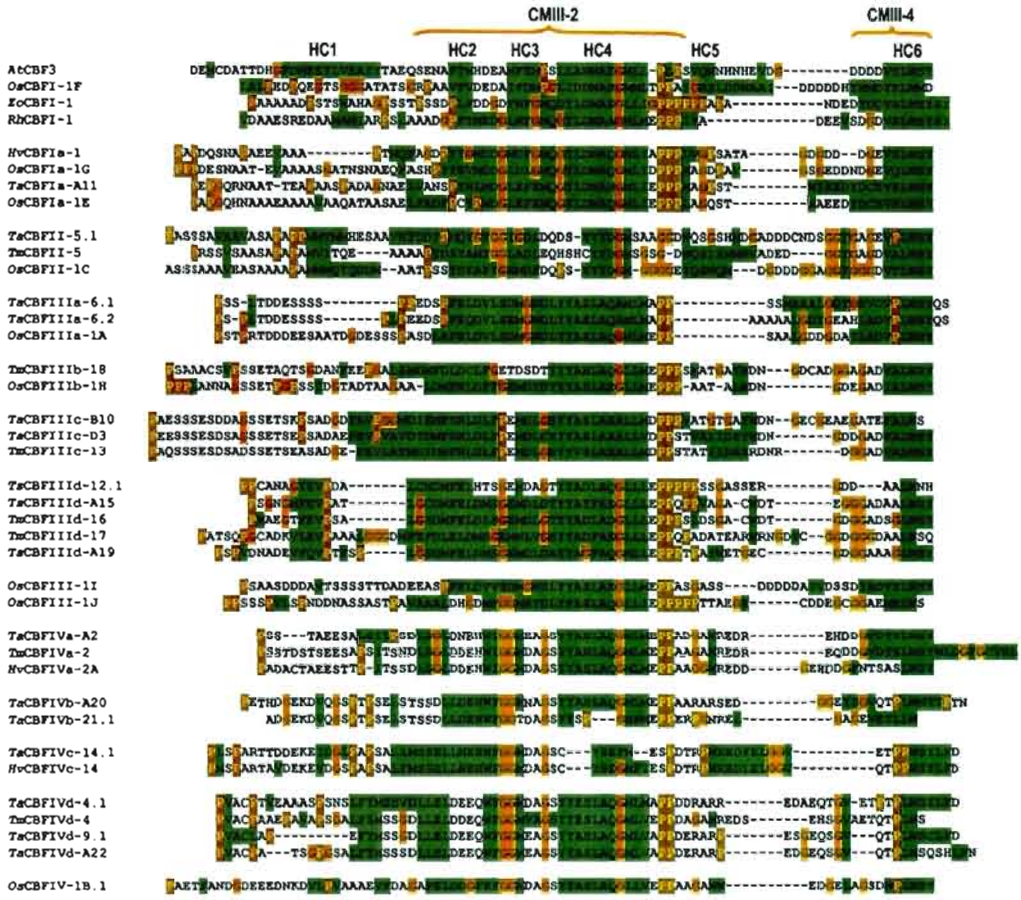
II-Figure 2: Comparative hydrophobic cluster analysis of the AP2 DNA-binding domain of wheat CBF protein groups.

The protein sequence of the AP2 DNA-binding domain and the regions CMIII-3 and CMIII-1 surrounding the AP2 domain of wheat, barley and rice CBF proteins were aligned using ClustalW and clusters of hydrophobic amino acids were highlighted in green, glycine in red and proline in dark red. For comparative purposes, this analysis was also done for the *Arabidopsis At*CBF3, *Oryza sativa Os*CBFI-1F, *Zingiber officinale Zo*CBFI-1 and *Rhapidophyllum hystrix Rh*CBFI-1 proteins. HC1 to HC8 identify hydrophobic clusters used to structurally define cereal CBF groups. Gaps (-) were introduced to maximize the alignment between all groups analyzed.


CMIII-3
AP2
CMIII-1
HC1
HC2
HC3
HC4
HC5
HC6
HC7
HC8
AtCBF3
OsCBFI-1F
ZoCBFI-1
RhCBFI-1
HvCBFIa-1
OsCBFIa-1G
TaCBFIa-A11
OsCBFIa-1E
TaCBFII-5.1
TmCBFII-5
OsCBFII-1C
TaCBFIIa-6.1
TaCBFIIa-6.2
OsCBFIIa-1A
TmCBFIIb-18
OsCBFIIb-1H
TaCBFIIc-D3
TaCBFIIc-B10
TmCRFIIc-13
TaCBFIId-12.1
TaCBFIId-A15
TmCBFIId-16
TmCBFIId-17
TaCBFIId-A19
OsCBFIII-1D
OsCBFIII-1I
OsCBFIII-1J
TaCBFIVa-A2
TmCBFIVa-2
HvCBFIVa-2A
TaCBFIVb-A20
TaCBFIVb-21.1
TaCBFIVc-14.1
HvCBFIVc-14
TaCBFIVd-4.1
TmCBFIVd-4
TaCBFIVd-9.1
TaCBFIVd-A22
OsCBFIV-1B.1

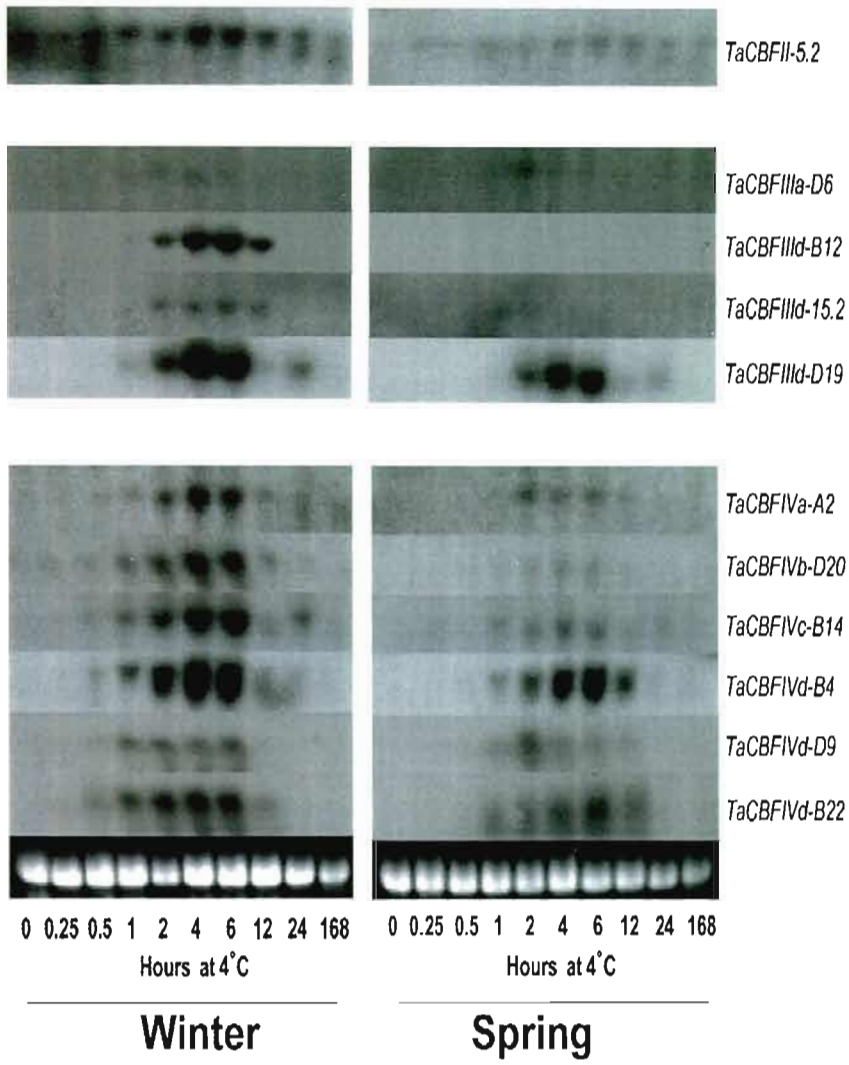
II-Figure 3: Comparative hydrophobic cluster analysis of the C-terminal region of wheat CBF proteins.

The regions surrounding motifs CMIII-2 and CMIII-4 of wheat CBF proteins were aligned using ClustalW and clusters of hydrophobic amino acids were highlighted in green, glycine in red and proline in dark red. For comparative purposes, this analysis was also done for the *Arabidopsis At*CBF3, *Oryza sativa Os*CBFI-1F, *Zingiber officinale Zo*CBFI-1 and *Rhapidophyllum hystrix Rh*CBFI-1 proteins. HC1 to HC6 identify hydrophobic clusters previously defined in *Arabidopsis* (Wang et al. 2005). Gaps (-) were introduced to maximize the alignment between members of a group, and to maximize the comparison of similar regions between groups


CMIII-2
CMIII-4
HC1
HC2
HC3
HC4
HC5
HC6
AtCBF3
OsCBFI-1F
EcCBFI-1
RbCBFI-1
HvCBFIa-1
OsCBFIa-1G
TaCBFIa-A11
OsCBFIa-1E
TaCBFII-5.1
TmCBFII-5
OsCBFII-1C
TaCBFIIIa-6.1
TaCBFIIIa-6.2
OsCBFIIIa-1A
TmCBFIIIb-18
OsCBFIIIb-1H
TaCBFIIIc-B10
TaCBFIIIc-D3
TmCBFIIIc-13
TaCBFIIId-12.1
TaCBFIIId-A15
TmCBFIIId-16
TmCBFIIId-17
TaCBFIIId-A19
OsCBFIII-1I
OsCBFIII-1J
TaCBFIVa-A2
TmCBFIVa-2
HvCBFIVa-2A
TaCBFIVb-A20
TaCBFIVb-21.1
TaCBFIVc-14.1
HvCBFIVc-14
TaCBFIVd-4.1
TmCBFIVd-4
TaCBFIVd-9.1
TaCBFIVd-A22
OsCBFIV-1B.1

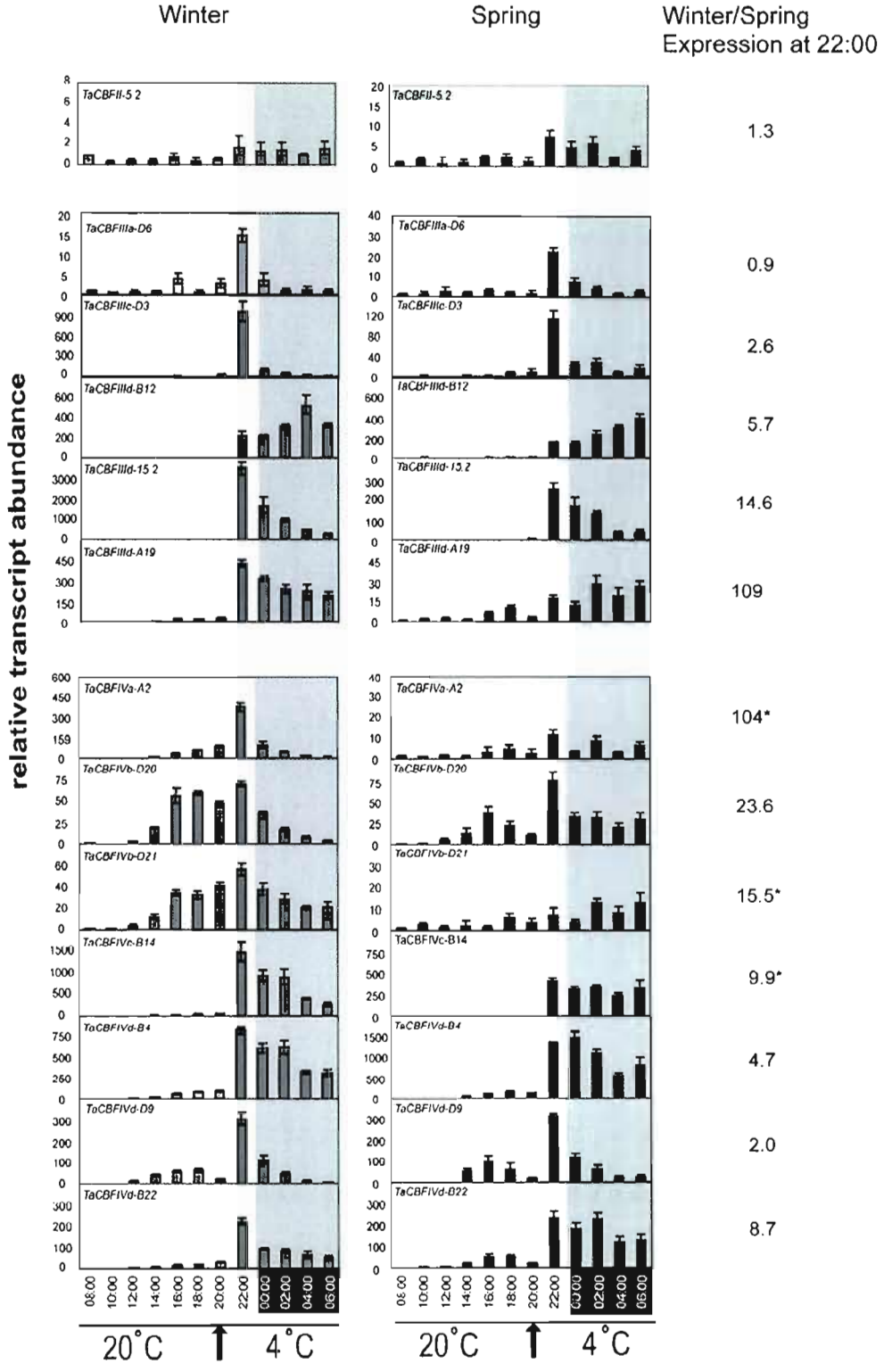
II-Figure 4: Northern analyses of 11 TaCBF transcripts during cold treatments of winter and spring wheats.

Total RNA was extracted from wheat leaves and analyses performed as described in the Materials and Methods. Hybridizations were done on a series of replicate blots with unique probes designed outside of the AP2 DNA-binding domain of 15 different *CBF* genes (Table S3), and the 11 *TaCBF* transcripts that displayed signals are shown. The cold treatment (4°C) was initiated 4 hours into the day phase and continued for the times indicated. The ethidium bromide-stained rRNA band that is shown is representative of each gel transferred to a membrane. The winter and spring cultivars were Norstar and Quantum, respectively

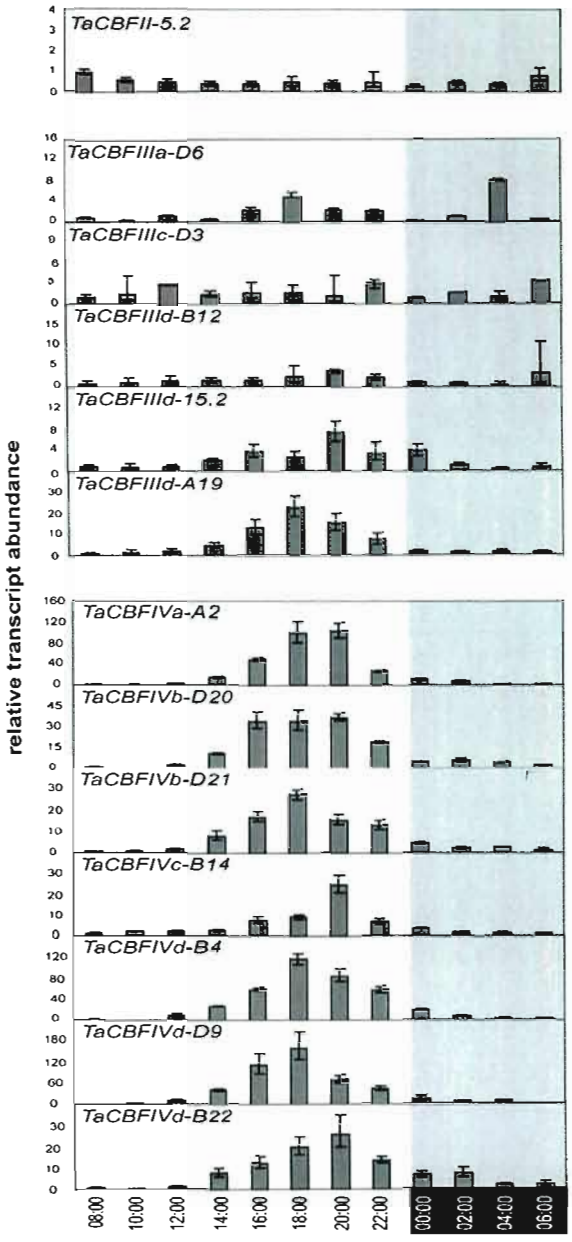

TaCBFII-5.2
TaCBFIIIa-D6
TaCBFIIId-B12
TaCBFIIId-15.2
TaCBFIIId-D19
TaCBFIVa-A2
TaCBFIVb-D20
TaCBFIVc-B14
TaCBFIVd-B4
TaCBFIVd-D9
TaCBFIVd-B22
0 0.25 0.5 1 2 4 6 12 24 168
Hours at 4°C
0 0.25 0.5 1 2 4 6 12 24 168
Hours at 4°C
Winter
Spring

II-Figure 5: Quantitative real-time PCR expression analysis of *TaCBF* genes in two wheat cultivars in response to low temperature.

Plants were germinated for 7 d at 20°C under a 16-h-day/8-h-night photoperiod. Beginning on day 8 at 8:00, plants were grown at 20°C for 12 h (from 08:00–20:00) then exposed at 4°C for 10 h (20:00-06:00). Gray areas represent the dark periods and arrows the start of the LT treatment. Leaf samples were harvested at each indicated time point, and RNA was extracted and analyzed by quantitative real-time PCR as described in Materials and Methods. Relative transcript abundance was calculated and normalized with respect to the 18S rRNA transcript level. Winter and spring panels were generated using the independent calibrators winter 08:00 and spring 08:00, respectively. This resulted in y-axis scales that are not directly comparable among cultivars but allow the direct visualization of the LT effect on gene expression within each cultivar. Error bars indicate the range of possible RQ values defined by the SE of the delta threshold cycles (Cts). The winter and spring cultivars were Norstar and Manitou, respectively. Winter/Spring Expression at 22:00 was determined by directly comparing the RQ values in both cultivars at the 22:00 time point. The RQ_{max} and RQ_{min} values did not exceed 25% deviation of the value shown except as indicated by (*) where the deviation ranged between 36 and 45%. It is noted that a different genome equivalent was assayed in this experiment (*TaCBFIIId-A19*) compared to Figure 4 (*TaCBFIIId-D19*) which prevents direct comparisons


Winter
Spring
Winter/Spring
Expression at 22:00
relative transcript abundance
TaCBFII-5 2
TaCBFIIIa-D6
TaCBFIIIc-D3
TaCBFIIId-B12
TaCBFIIId-15 2
TaCBFIIId-A19
TaCBFIVa-A2
TaCBFIVb-D20
TaCBFIVb-D21
TaCBFIVc-B14
TaCBFIVd-B4
TaCBFIVd-D9
TaCBFIVd-B22
1.3
0.9
2.6
5.7
14.6
109
104*
23.6
15.5*
9.9*
4.7
2.0
8.7
08:00
10:00
12:00
14:00
16:00
18:00
20:00
22:00
00:00
02:00
04:00
06:00
20°C
4°C

II-Figure 6: Quantitative real-time PCR expression analysis of *TaCBF* genes in Norstar in response to a diurnal cycle. Plants were germinated for 7 d at 20°C under a 16-h-day/8-h-night photoperiod. Beginning on day 8 at 8:00, plants were grown at 20°C for 22 h (from 08:00–06:00). Gray areas represent the dark periods. Leaf samples were harvested at each time point, and RNA was extracted and analyzed by quantitative real-time PCR as described in Materials and Methods. Relative transcript abundance was calculated and normalized with respect to the 18S rRNA transcript level and the calibrator time point (08:00). Error bars indicate the range of possible RQ values defined by the SE of the delta threshold cycles (Cts)


relative transcript abundance
TaCBFII-5.2
TaCBFIIIa-D6
TaCBFIIIc-D3
TaCBFIIId-B12
TaCBFIIId-15.2
TaCBFIIId-A19
TaCBFIVa-A2
TaCBFIVb-D20
TaCBFIVb-D21
TaCBFIVc-B14
TaCBFIVd-B4
TaCBFIVd-D9
TaCBFIVd-B22
08:00
10:00
12:00
14:00
16:00
18:00
20:00
22:00
00:00
02:00
04:00
06:00

II- Supplemental DATA

II- Table S1: Summary of the different wheat cDNA libraries used in this project

Library	Growth conditions[1]	Tissues	Methyl dCTP	Full length[2]
Library 2	Control plants; Plants cold acclimated for 1, 23 and 53 days	leaves and crowns		
Library 3	Control plants; Plants cold acclimated for 1, 23 and 53 days; Plants salt stressed for 0.5, 3 and 6 hours	roots	yes	
Library 4	Plants dehydrated on the bench (4 time points) and in a growth chamber (4 time points)	leaves and crowns	yes	
Library 5	Various vernalization and developmental stages through spike formation	crowns and flowers	yes	
Library 6	Control plants; Plants cold acclimated for short time points (1, 3 and 6 hours) under light or dark conditions	leaves and crowns	yes	yes

[1]Growth conditions are described in more detail elsewhere (Houde et al. 2006).
[2]The full length library was used to isolate *TaCBF* genes and identify possible transcription start sites.

II- Table S2: Gene specific primers used to amplify *TaCBF* genes

Gene Name	Clone Name	Primer Name	Direction	Primer Sequence
TaCBFIa-A11	BQ744365F1L8_153	BQ744365F1	Forward	5'-CACCTCGAACGGACTGATCACTGATGGA-3'
		BQ744365F1L8-153R1	Reverse	5'-CCACATTGCCATGACATTTTCGTCAGAATAGCAG-3'
TaCBFII-5.1	L4B201_C08	L4B201_C08F1	Forward	5'-CACACAAACAACCACCCAGCATCACCTCAC-3'
		L4B201_C08R1	Reverse	5'-GGAGGGAAACAAACAGGCTGATGCAATTCG-3'
TaCBFII-5.2	C21R1L8_A	C21R1	Reverse	5'-CCCAACACATGTTACAAAATTTCCAGCCAACAA-3'
TaCBFII-5.3	C22R1L3_A	C22R1	Reverse	5'-TGAAGCTCCCGTTCCCTCTCTGCTTCTTCTT-3'
TaCBFIIIa-6.2	C25R1L1_48	C25R1	Reverse	5'-CCCCAATTCGCACCATCACATTACCAGA-3'
TaCBFIIIc-3.1	C16R1L7_AF1_L8_41	C16R1L7_AF1	Forward	5'- CCATGGACATGGGCCTTGAGGTCTCG-3'
TaCBFIIIc-3.2	C16R1L7_AR1_L1_2	C16R1	Reverse	5'-CCCACAAGGGGAAAACTTTGTAACAGGGGAAA-3'
TaCBFIIId-12.1	L6B013_O03R1L1_4	L6B013_O03R1L1_4F1	Forward	5'-GTCATAGCCACCAAGGCAGCAGCCATCAC-3'
		L6B013_O03R1	Reverse	5'-CATCATGGGTTCACACCGTGCACTCGTA-3'
TaCBFIIId-A15	L6B002_G24R2L1_13	L6B002_G24F1	Forward	5'-GCTCAAGCTCAGCCTGCTCTCACACTCCAT-3'
		L6B002_G24R2	Reverse	5'-TCCACAATTTTATTATGCTGACAGGGAACAGCTTCG-3'
TaCBFIIId-15.2	L6B002_G24F1L1_49	L6B002_G24F1	Forward	5'-GCTCAAGCTCAGCCTGCTCTCACACTCCAT-3'
		L6B002_G24R2	Reverse	5'-TCCACAATTTTATTATGCTGACAGGGAACAGCTTCG-3'
TaCBFIVa-2.2	D992F2L6_2	D992F2	Forward	5'-CGGTGTGGCCGGAGGAGCAGGAGTACC-3'
TaCBFIVa-2.3	GI25164819R1L1_7	GI25164819R1	Reverse	5'-ACGGAGCATCCCCTTGCGCCAAGCTC-3'
TaCBFIVb-A20	Clone9F1L7_5	Clone9F1	Forward	5'-CTGCAACTCTCAACGCAGCGACTTTCCAC-3'
		L3C116_P09-23R2	Reverse	5'-ATGCCGCCAAACCAGTGCTCGTCCAAC-3'
TaCBFIVb-21.1	C133R1L1_2F1L7_7	C133R1L1-2F1	Forward	5'-AGCTGCAATCCATCGACCCTCACCACAAC-3'
		C133R1L1-4R1	Reverse	5'-AGCAAGTCGCTGGACGTGGACAGCTC-3'
TaCBFIVb-D21	C133R1L1_4	C133R1	Reverse	5'-CCGGAGGCTTTCAATATGAACGAACACACA-3'
TaCBFIVd-A22	CA677056F1L1_143	CA677056F1	Forward	5'-CGGAAACACAGCAACCTGCTCAAACGTCTC-3'
		CA677056F1L1-143R1	Reverse	5'-TGCTAAATTAGTTGAACAGGTGGCTCTGGCTCCAT-3'
TaCBFIVd-B22	L6B020_O14F1L1_72	L6B020_O14F1	Forward	5'-CAAACACCGCCGCTGAAACCTCCAGTTC-3'
		L6B020_O14R2	Reverse	5'-TTTAATCCATTTCGGCACAAGTAATTCCGGACAG-3'
TaCBFIVd-D22	L6B020_O14F1L1_68	L6B020-O14F1	Forward	5'-CAAACACCGCCGCTGAAACCTCCAGTTC-3'
		GI25237880R1	Reverse	5'-TGCAATCCCCTTCCTCACAAACACTCAAACC-3'

Forward gene specific primers were used in conjunction with M13F and reverse gene specific primers were used with M13R to amplify 3' or 5' regions of *TaCBF* cDNAs, respectively. When both regions were amplified, overlapping sections (100% identity) were merged to present the longest sequence.

II- Table S3: Primer pairs used to amplify the probes for Northern blot experiments

Primer Name	Direction	Primer Sequence	Amplicon size (pb)
*Ta*CBF1a-A11 F	Forward	5'-TAGAATTCGAAATGCAGGGCTATC-3'	291
*Ta*CBF1a-A11 R	Reverse	5'-CACTAATTGTTTGTTGACCCACAT-3'	
*Ta*CBFII-5.2 F	Forward	5'-CACACAAACACCCAACCAACATCAC-3'	320
*Ta*CBFII-5.2 R	Reverse	5'-AGCCAGATGCGGGACTTCTTGTT-3'	
*Ta*CBFIIIa-D6 F	Forward	5'-ACCGACGACGGGTCGTCCTCCT-3'	305
*Ta*CBFIIIa-D6 R	Reverse	5'-ACTGCAAGCGGTGACCCTAAAGCA-3'	
*Ta*CBFIIIc-D3 F	Forward	5'-ACTTGTTCCCGGAAATGGACCTG-3'	267
*Ta*CBFIIIc-D3 R	Reverse	5'-CCAAAATAGGAAAACTTTGTAACAGGGG-3'	
*Ta*CBFIIIc-B10 F	Forward	5'-GAGGCGCTGCTCATGGAC-3'	150
*Ta*CBFIIIc-B10 R	Reverse	5'-AAAGACGCTACAGAGTCAAAACAAA-3'	
*Ta*CBFIIId-B12 F	Forward	5'-ACACGTCCGGCGAAATGGA-3'	272
*Ta*CBFIIId-B12 R	Reverse	5'-CCAAATGTCCCGCGATTTTATTAT-3'	
*Ta*CBFIIId-15.2 F	Forward	5'-TACCTTCTGGCGACGGAATGTTT-3'	224
*Ta*CBFIIId-15.2 R	Reverse	5'-GCAGCTGGCTGGAGTGTTTTAGTA-3'	
*Ta*CBFIIId-D19 F	Forward	5'-CCACACTCGTCTACCCAACA-3'	266
*Ta*CBFIIId-D19 R	Reverse	5'-CTCCTTGAACTTGGTGCGC-3'	
*Ta*CBFIVa-A2 F	Forward	5'-TCGTACTACGCGAGCTTGGCGC-3'	267
*Ta*CBFIVa-A2 R	Reverse	5'-TGTGCCCTTCCGGGAGTAGAAACC-3'	
*Ta*CBFIVb-D20 F	Forward	5'-CCGCCAGAACGTGGAGCGAG-3'	244
*Ta*CBFIVb-D20 R	Reverse	5'-CTCCTGATGCTTACTCTGTATCTTCCCATA-3'	
*Ta*CBFIVb-21.1 F	Forward	5'-AGCTGTCCACGTCCAGCGATTT-3'	305
*Ta*CBFIVb-21.1 R	Reverse	5'-TCACAGAACACGCAAGTGCAAA-3'	
*Ta*CBFIVc-B14 F	Forward	5'-GATGCCGGATCGTTCTACTC-3'	110
*Ta*CBFIVc-B14 R	Reverse	5'-AAGTAGCTCCATGGCGGTGT-3'	
*Ta*CBFIVd-B4 F	Forward	5'-TCGACGAGGAGCACTGGTTT-3'	305
*Ta*CBFIVd-B4 R	Reverse	5'-AGGGTCCAATTCTCACAAAAGTAG-3'	
*Ta*CBFIVd-D9 F	Forward	5'-GCGGAAGCAGATTATTCCGGTCGC-3'	376
*Ta*CBFIVd-D9 R	Reverse	5'-CGTCTGGACGCCGCTCTGCT-3'	
*Ta*CBFIVd-B22 F	Forward	5'-CTCAGTTCTCACCCACCCAAACA-3'	210
*Ta*CBFIVd-B22 R	Reverse	5'-GTGGCGCGTCTCATGTACCTT-3'	

II- Table S4: List of forward, reverse and fluorescent TaqMan-MGB probes used in quantitative real-time PCR analyses of *TaCBF* genes

Primer Name	Type	Primer Sequence	Amplicon size (bp)	Position
*Ta*CBFIa-A11 F	Forward	5'-AACCAAGCCAGGCACACAAG-3'	62	9-28
*Ta*CBFIa-A11 R	Reverse	5'-CTGTACGCCCACTCCATCAGT-3'		70-50
*Ta*CBFIa-A11 P	Probe	5'-CACCTCGAACGGACTG-3'		31-46
*Ta*CBFII-5.2 F	Forward	5'-AACATCACCTCACTCACCAGTCA-3'	103	18-40
*Ta*CBFII-5.2 R	Reverse	5'-GTACTGGTCCATGGTGTTGCA-3'		120-100
*Ta*CBFII-5.2 P	Probe	5'-ATCGGCACCGGCTA-3'		74-87
*Ta*CBFIIIa-D6 F	Forward	5'-CCAAGCCGGCCAAGAAA-3'	68	49-65
*Ta*CBFIIIa-D6 R	Reverse	5'-TCGGACACATCTTCTTCCTTGAAT-3'		116-93
*Ta*CBFIIIa-D6 P	Probe	5'-CGAGCAAAACCTCTC-3'		75-89
*Ta*CBFIIIc-D3 F	Forward	5'-TGGACCCGCCGTCGA-3'	70	757-771
*Ta*CBFIIIc-D3 R	Reverse	5'-GCGACATCAGCTCCGTCGT-3'		826-808
*Ta*CBFIIIc-D3 P	Probe	5'-CCATCATCGACTCGTACT-3'		778-795
*Ta*CBFIIIc-B10 F	Forward	5'-TCCGCCGATGGAGACTTC-3'	64	595-612
*Ta*CBFIIIc-B10 R	Reverse	5'-AGTCAAGCCTACTGAACATTTCGA-3'		658-635
*Ta*CBFIIIc-B10 P	Probe	5'-CATTCCGCCCGGCA-3'		630-617
*Ta*CBFIIId-B12 F	Forward	5'-AGTACTCCTACCCATAGCCACCAA-3'	87	5-28
*Ta*CBFIIId-B12 R	Reverse	5'-CCATTGCCGCGGTACGTA-3'		91-74
*Ta*CBFIIId-B12 P	Probe	5'-CCTCCAGTCAACTAGT-3'		48-63
*Ta*CBFIIId-15.2 F	Forward	5'-GATGACGGAAGCTGTAGCCAAT-3'	102	551-572
*Ta*CBFIIId-15.2 R	Reverse	5'-AAACATTCCGTCGCCAGAAG-3'		652-633
*Ta*CBFIIId-15.2 P	Probe	5'-TACCGGCTCGTCGTG-3'		595-609
*Ta*CBFIIId-A19 F	Forward	5'-GCCTCTGGAGCTACTGATGTCTG-3'	68	872-894
*Ta*CBFIIId-A19 R	Reverse	5'-CAATCGGGATAGCAAAAATCCTC-3'		939-917
*Ta*CBFIIId-A19 P	Probe	5'-ACCTCCAGTGGGTTC-3'		897-911
*Ta*CBFIIId-B19 F	Forward	5'-GCCGGCCTCTGGAGCTA-3'	73	834-850
*Ta*CBFIIId-B19 R	Reverse	5'-TCCAATCGGGATAACAAAATCC-3'		906-885
*Ta*CBFIIId-B19 P	Probe	5'-AACCTTCAGTGGGTTCA-3'		862-878
*Ta*CBFIIId-D19 F	Forward	5'-GCCGGCCTCTGGAGCTA-3'	75	850-866
*Ta*CBFIIId-D19 R	Reverse	5'-GGTCCAATCGGGATAACAAAATC-3'		924-902
*Ta*CBFIIId-D19 P	Probe	5'-CACTGCAGGTTACAGACA-3'		888-871
*Ta*CBFIVa-A2 F	Forward	5'-GCCGGAAGCCGGAGTTA-3'	61	259-275
*Ta*CBFIVa-A2 R	Reverse	5'-CCGCCATCTGGGCACT-3'		319-304
*Ta*CBFIVa-A2 P	Probe	5'-TCTGGCTTGGCACCT-3'		284-298
*Ta*CBFIVb-A20 F	Forward	5'-AGCCATGTACACTTTTAGAAACTACTCCT-3'	89	730-758
*Ta*CBFIVb-A20 R	Reverse	5'-TGCTCCTGATGCTTACTCTGTATCTT-3'		818-793
*Ta*CBFIVb-A20 P	Probe	5'-TGTTCGTCCAAATATG-3'		775-790
*Ta*CBFIVb-B20 F	Forward	5'-TCGTCCCGCTTCAGCA-3'	103	390-405
*Ta*CBFIVb-B20 R	Reverse	5'-CGTGCGTCTCCGACGTG-3'		492-476
*Ta*CBFIVb-B20 P	Probe	5'-CCATCGTGGAGTTCC-3'		439-453
*Ta*CBFIVb-D20 F	Forward	5'-GAGGATGGCGGCGAATAC-3'	90	668-685
*Ta*CBFIVb-D20 R	Reverse	5'-GAGTAGTTTCCAAAACTGTACATGGCTTA-3'		757-729
*Ta*CBFIVb-D20 P	Probe	5'-AGCGCCGTCTACAC-3'		686-699
*Ta*CBFIVb-D21 F	Forward	5'-TCATCGGAGCCCCAATCA-3'	56	312-329
*Ta*CBFIVb-D21 R	Reverse	5'-CGTGCGTTAAGCCGTTGTG-3'		367-349
*Ta*CBFIVb-D21 P	Probe	5'-TCGAACCTCGCCAGTC-3'		331-346
*Ta*CBFIVc-B14 F	Forward	5'-GACACCGCCATGGAGCTACT-3'	93	720-739
*Ta*CBFIVc-B14 R	Reverse	5'-CCTCCCATATTTGGTGGAACAG-3'		812-791
*Ta*CBFIVc-B14 P	Probe	5'-TGTTTGTATAGATAGTTTCTCTC-3		766-788

II- Table S4: continued

*Ta*CBF1Vd-B4 F	Forward	5'-CCAGCTGCAGCAGAATATTCCT-3'	93	457-478
*Ta*CBF1Vd-B4 R	Reverse	5'-TCGCCGGACGACATGTAGA-3'		549-531
*Ta*CBF1Vd-B4 P	Probe	5'-CTGTCGGTCGGCTC-3'		500-513
*Ta*CBF1Vd-B9 F	Forward	5'-AACACAGCCGCTGATTCCA-3'	61	4-22
*Ta*CBF1Vd-B9 R	Reverse	5'-CTAGCGGAGATGCTCGTGAGA-3'		64-44
*Ta*CBF1Vd-B9 P	Probe	5'-ACTACTACCACTCCACAC-3'		25-42
*Ta*CBF1Vd-D9 F	Forward	5'-GATCGATCAAAACCTCTCAACACA-3'	72	19-42
*Ta*CBF1Vd-D9 R	Reverse	5'-AGACGCTCGTGGGAGGTG-3'		90-73
*Ta*CBF1Vd-D9 P	Probe	5'-CAGTACTCCTGCTCCA-3'		57-72
*Ta*CBF1Vd-A22 F	Forward	5'-GCGTCCAGACGCCGTTAT-3'	77	879-896
*Ta*CBF1Vd-A22 R	Reverse	5'-ATGCTATCTATGGTACAGCTTTACACTGC-3'		955-927
*Ta*CBF1Vd-A22 P	Probe	5'-CCACCTGTTCAACTAA-3'		907-922
*Ta*CBF1Vd-B22 F	Forward	5'-CCGCTATGGAGCCACTTGTT-3'	66	927-946
*Ta*CBF1Vd-B22 R	Reverse	5'-TAATTCCGGACAGAGGGAGTACA-3'		992-970
*Ta*CBF1Vd-B22 P	Probe	5'-CACCTAATCTAGCAGTGTAAA-3'		967-947
*Ta*CBF1Vd-D22 F	Forward	5'-GAGCCAGAGCCACTTGTTCA-3'	105	899-918
*Ta*CBF1Vd-D22 R	Reverse	5'-ACAAGATGCTACTGTGTTTCTCTCCAA-3'		1003-977
*Ta*CBF1Vd-D22 P	Probe	5'-CCATAGTCCATAGATAGTTG-3'		944-963

III-Table S5: List of nucleotide sequences encoding the AP2 DNA binding region of monocot *CBF*s used for the phylogenetic analysis

>OsCBFI-1F

CCGAAGCGGCGGGCGGGGCGGAAGAAATTCCGGGAGACGCGGCACCCGGTGTACCGCGGCGTGCGCGC
GCGGGCGGGGGGGAGCAGGTGGGTGTGCGAGGTGCGCGAGCCGCAGGCGCAGGCGCGCATCTGGCTCG
GCACCTACCCGACGCCGGAGATGGCGGCGCGCGCGCACGACGTCGCGGCCATCGCCCTCCGCGGCGAG
CGCGGCGCCGAGCTCAACTTCCCGGACTCCCCCTCCACG

>TaCBFIa-A11

CCCAAGCGGCGGGCGGGGAGGACCAAGGTCAGGGAGACGAGGCACCCGGTGTACAAGGGGGTGCGCAG
CAGGAACCCCGGCCGGTGGGTCTGCGAGGTGCGCGAGCCGCACGGGAAGCAGCGGCTGTGGCTCGGCA
CCTTCGACACCGCCGAGATGGCGGCGCGCGCGCACGACGTCGCCGCCCTCGCGCTCCGCGGCCGCGCC
GCGTGCCTCAACTTCGCGGACTCGCCGCGCACG

>HvCBFIa-11

CCCAAGCGGCGGGCCGGGAGGACCAAGGTCAGGGAGACGAGGCACCCGGTGTACAAGGGGGTGCGCAG
CAGGAACCCCGGGCGGTGGGTCTGCGAGGTGCGCGAGCCGCAAGGGAAGCAGAGGCTATGGCTCGGCA
CCTTCGACACCGCCGAGATGGCGGCGCGCGCGCACGACGTCGCCGCCATGGCGCTCCGCGGCCGCGCC
GCGTGCCTCAACTTCGCGGACTCGCCGCGGAGG

>SoCBFIa-11

CCCAAGCGGCGCGCGGGGCGGACCAAGTTCAAGGAGACGAGGCACCCGGTGTACAAGGGCGTGCGCAG
CCGGAACCCCGGGCGGTGGGTCTGCGAGGTGCGGGAGCCGCACGGCAGGCAGCGGATCTGGCTCGGCA
CCTTCGAGACCGCCGAGATGGCGGCGCGCGCGCACGACGTCGCCGCGCTCGCGCTGCGTGGCCGCGCC
GCGTGCCTCAACTTCGCGGACTCGCCGCGCAGG

>PvCBFIa-11

CCGAAGCGGCGCGCGGGGCGGACCAAGTTCAAGGAGACGCGGCACCCGGTGTACAAGGGCGTGCGCAG
CCGGAACCCCGGGCGCTGGGTCTGCGAGCTGCGGGAGCCGCACGGCAGGCACCGCATCTGGCTCGGCA
CCTTCGAGACCGCCGAGATGGCGGCGCGCGCGCACGACGTCGCCGCGCTGGCGCTGCGCGGCAGGGCC
GCCTGCCTCAACTTCGCCGACTCGCCGCGCAGG

>OsCBFIa-1E

CCCAAGCGGCGGGCGGGGAGGACCAAGTTCAAGGAGACGAGGCACCCGGTGTACAAGGGCGTGCGCAG
CAGGAACCCCGGGAGGTGGGTCTGCGAGGTGCGCGAGCCGCACGGCAAGCAGAGGATCTGGCTCGGCA
CCTTCGAGACCGCCGAGATGGCGGCGCGCGCGCACGACGTCGCGGCGATGGCGCTGCGCGGCCGCGCC
GCCTGCCTCAACTTCGCCGACTCGCCGCGCAGG

>OsCBFIa-1G

CCCAAGCGGCGAGCGGGGCGGACCAAGTTCAAGGAGACGCGGCACCCCGTGTTCAAGGGCGTGCGCCG
GAGGAACCCCGGGAGGTGGGTGTGCGAGGTGCGCGAGCCGCACGGCAAGCAGCGGATATGGCTCGGGA

CGTTCGAGACAGCAGAGATGGCGGCGCGCGCGCACGACGTCGCCGCGCTCGCGCTCCGCGGCCGCGCC
GCCTGCCTCAACTTCGCCGACTCGCCGAGGCGC

>ZoCBFI-1

CCGAAGCGCCGGGCCGGCCGCACCAAGTTCCGGGAGACGCGCCACCCGGTGTTCAAAGGCGTCCGGCG
CCGGAACGGCGACAAGTGGGTGTGCGAGGTCCGGGAGCCCAACAAGAAGTCCCGAATTTGGCTCGGCA
CCTTCCCCACCGCCGAGATGGCTGCCCGCGCGCACGACGTCGCCGCCATCGCCCTCCGCGGCCGCGCC
GCCTGCCTCAACTTCGCCGACTCCGCCTGGCTC

>ZoCBFI-2

CCGAAGCGGCCGGCGGGGCGGACCAAATTCAGGGAGACGCGGCACCCGGTGTACAAGGGCGTCCGGCG
GCGGGGAACGGCGGGGCGGTGGGTGTGCGAGGTGCGGGAACCCAACAAGAAGTCTAGGATCTGGCTAG
GGACCTTCCCAACGGCAGATATGGCGGCGCGCGCCCACGACGCCGCCGCCATGATGCTCCGCGGCCGC
GCAGCCTGCCTCAATTTCGCCGACTCCGCCTGGCTC

>SpCBFI-1

CCGAAGCGGCGGGCTGGGCGGACCAAGTTCCGCGAGACGCGACACCCCGTCTACAAAGGCGTCCGGCG
GCGAAACGCCGATAAGTGGGTCTGCGAAGTGCGCGAGCCCAACAAGAAGTCCAGGATTTGGCTCGGGA
CGTTTCCGACCGCCGAGATGGCGGCACGGGCGCACGACGTGGCTGCCATGGCCCTCCGCGGCCGGTCG
GCCTGCCTCAACTTTGCTGATTCCGCGTGGCTC

>RhCBFI-1

CCGAAGCGGCGGGCTGGGCGGACCAAGTTCCGCGAGACGCGACACCCGGTTTACAAGGGCGTCCGGCG
CCGCAACGCCGATAAGTGGGTCTGCGAAGTGCGGGAGCCCAACAAAAAGTCGAGGATTTGGCTCGGGA
CGTTTCCGACCGCCGAGATGGCTGCACGGGCGCACGACGTAGCTGCCATGGCCCTCCGCGGCCGATCA
GCCTGCCTCAACTTTGCTGATTCCGCGTGGCTC

>RhCBFI-2

CCGAAGAGGCGGGCCGGGCGGACCAAGTTTCGCGAGACGCGGCACCCGGTCTACAAGGGCGTCCGCCG
CCGGAACGCCGACAAGTGGGTCTGCGAGGTGCGTGAGCCCAACAAGAAGTCGAGGATTTGGCTCGGGA
CGTTTCCCACCGCCGAGATGGCTGCACGGGCGCACGACGTGGCTGCCATGGCCCTCCGCGGCCGGTCG
GCCTGCCTCAACTTTGCCGACTCCGCTTGGCTC

>DlCBFI-1

CCGAAGAGGCGGGCTGGAAGGACCAAATTCCGCGAGACGCGACACCCGGTCTACAAGGGCGTCCGACG
CCGCAACGCCGATAAGTGGGTGTGCGAGGTGCGTGAGCCCAACAAAAAGTCGAGGATTTGGCTCGGGA
CGTTTCCGACCGCTGAGATGGCGGCACGGGCGCACGACGTGGCTGCCATGGCCCTCCGCGGTCGGTCA
GCATGCCTCAACTTTGCTGATTCGGCGTGGCTC

>HvCBFIa-1

CCGAAGCGGCGCGCGGGGCGGACCAAGTTCAAGGAGACGCGGCACCCGGTCTACAAGGGCGTGCGCCG
GAGGAACCCCGGGAGGTGGGTCTGCGAGGTGCGGGAGCCGCACAGCAAGCAGAGGATATGGCTCGGCA
CGTTCGAGACCGCAGAGATGGCGGCGCGCGCGCACGACGTGGCCGCGCTGGCGCTGCGCGGCCGCGCC
GCCTGCCTCAACTTCGCCGACTCTCCTCGTCGG

>BdCBFIa-1

CCGAAGCGGCGTGCGGGGCGGACCAAGTTCAAGGAGACGCGGCACCCGGTGCTCAAGGGCGTGCGCCG
GAGGAACCCCGGGAGGTGGGTCTGCGAGGTGCGGGAGCCCCACAGCAAGCAGAGGATATGGCTCGGGA
CATTCGAGACCGCCACGCTCAACTTCGCCGACTCGCCCCGCAGG

>ZmCBFII-5

CCCAAGCGGCCGGCGGGGCGGACCAAGTTCCGGGAGACGCGGCACCCCGTGTACCGCGGCGTGCGGCG
GCGCGGGCCCGCGGGGCGCTGGGTGTGCGAGGTCCGCGAGCCCAACAAGAAGTCGCGCATCTGGCTCG
GCACCTTCGCCACCCCCGAGGCCGCCGCGCGCGCGCACGACGTGGCCGCGCTGGCCCTGCGGGGCCGC
GCCGCGTGCCTCAACTTCGCCGACTCGGCGCGCCTG

>BdCBFII-5

CCGAAGCGGCCGGCGGGGCGGACAAAGTTCAGGGAGACGCGGCACCCGGTGTACCGCGGCGTGCGGCG
GCGCGGCGCGGCCGGGCGGTGGGTGTGCGAGGTCAGGGAGCCCAACAAGAAGTCCCGCATCTGGCTCG
GCACGTTCGCGTCCGCCGAGGCCGCCGCGCGCGCCCACGACGTCGCCGCCCTGGCGCTCCGTGGCCGC
GCCGCGTGCCTCAACTTCGCCGACTCGGCCGCGCTG

>PvCBFII-5

CCCAAGCGCCCCGCGGGGCGGACCAAGTTCCGGGAGACCCGCCACCCGGTGTACCGCGGCGTGCGGCG
GCGCGGCCCCGCGGGGCGCTGGGTCTGCGAGGTCCGGGAGCCCAACAAGAAGTCCCGCATCTGGCTCG
GCACCTTCGCCACCGCCGAGGCCGCCGCGCGCGCGCACGACGTCGCCGCGCTCGCGCTCCGGGGCCGC
GCGGCCTGCCTCAACTTCGCCGACTCGGCGCGCCTG

>SbCBFII-5

CCGAAGCGCCCCGCGGGTCGGACCAAGTTCCGGGAGACGCGTCACCCGGTGTACCGCGGCGTGCGGCG
GCGCGGGCCCGCGGGGCGGTGGGTGTGCGAGGTCCGGGAGCCCAACAAGAAGTCCCGCATCTGGCTCG
GCACCTTCGCCACCGCCGAGGCCGCCGCGCGCGCGCACGACGTCGCCGCGCTCGCGCTCCGCGGCCGA
GCCGCCTGCCTCAACTTCGCCGACTCCGCGCGCCTG

>SoCBFII-5

CCGAAGCGCCCAGCGGGGCGGACCAAGTTCCGGGAGACGCGTCACCCGGTGTACCGCGGCGTGCGGCG
GCGCGGGCCCGCGGGGCGGTGGGTGTGCGAGGTCCGGGAGCCCAACAAGAAGTCCCGCATCTGGCTCG
GCACCTTCGCCACCGCCGAGGCCGCCGCGCGCGCGCACGACGTCGCCGCGCTCGCGCTCCGCGGGCGC
GCCGCCTGCCTCAACTTCGCCGACTCCGCGCGCCTG

>TmCBFII-5

CCGAAGCGGCCGGCGGGGCGGACCAAGTTCCGGGAGACGCGCCACCCGGTGTACCGCGGCGTGCGCCG
GCGCGGCGCCGCGGGGCGCTGGGTGTGCGAGGTGCGCCAGCCCAACAACAAGTCCCGCATCTGGCTCG
GCACCTTCGCCAGCCCCGAGGCCGCCGCGCGCGCCCACGACGTCGCCGCGCTCGCGCTCCGGGGCCGC
GCCGCCTGCCTCAACTTCGCCGACTCGGCCGCGCTG

>OsCBFII-1C

CCGAAGCGGCCGGCGGGGAGGACCAAGTTCAGGGAGACGAGGCACCCGGTGTACCGCGGCGTGCGGCG
GCGGGGGCCCGCGGGGCGGTGGGTGTGCGAGGTCAGGGAGCCCAACAAGAAGTCCCGCATCTGGCTCG

GCACCTTCGCCACCGCCGAGGCCGCCGCGCGCGCCCACGACGTCGCCGCGCTCGCCCTCCGCGGCCGC
GGCGCGTGCCTCAACTTCGCCGACTCGGCCCGCCTC

>HvCBFII-5

CCCAAGCGGCCGGCGGGGCGGACCAAGTTCAGGGAGACGCGGCACCCGGTGTACCGGGGCGTGCGGCG
GCGCGGCGCCGCAGGGCGGTGGGTCTGCGAGGTGCGCGAGCCCAACAAGAAGTCCCGCATCTGGCTCG
GCACCTTTGCCACGCCCGAGGCCGCCGCGCGCGCCCACGACGTCGCCGCGCTCGCGCTCCGGGGCCGC
GCCGCCTGCCTCAACTTCGCCGACTCCGCGGCCCTG

>TaCBFII-5.1

CCGAAGCGGCCGGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCGGTGTACCGGGGCGTGCGCCG
GCGCGGCGCGGCCGGGCGGTGGGTCTGCGAGGTGCGCGAGCCCAACAAGAAGTCCCGCATCTGGCTCG
GCACCTTCGCCTCCCCCGAGGCCGCGGCCCGCGCCCACGACGTTGCCGCGCTGGCCCTGCGCGGCCGC
GCCGCGTGCCTCAACTTCGCCGACTCGGCGGCCCTG

>HvCBFIIIa-6

GCGAAGCGGCCGGCGGGGCGGACCAAGTTCAGGGAGACGCGCCACCCGGTGTACCGCGGCGTGCGGCG
CAGGGGCAATGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCCGGCAGGCGCGGGAGCAGGCTCTGGC
TCGGCACCTTCGACACCGCCGAGGCCGCCGCGCGCGCGAACGACGCCGCCATGCTCGCGCTCGCCGCC
GGGGGCGCGGGCTGCCTCAACTTCGCCGACTCCGCCGAGCTA

>OsCBFIIIa-1A

CCGAAGAGGCCGGCGGGGCGGACCAAGTTCAGGGAGACGAGGCACCCGGTGTTCCGCGGCGTGCGGCG
GAGGGGCAATGCCGGGAGGTGGGTGTGCGAGGTGCGGGTGCCCGGGCGGCGCGGCTGCAGGCTCTGGC
TCGGCACGTTCGACACCGCCGAGGGCGCGGCGCGCGCGCACGACGCCGCCATGCTCGCCATCAACGCC
GGCGGCGGCGGCGGCGGGGGAGCATGCTGCCTCAACTTCGCCGACTCCGCGTGGCTC

>TaCBFIIIa-6.1

GCGAAGCGGCCGGCGGGGCGGACCAAGTTCAGGGAGACGCGGCACCCGGTATACCGCGGCGTGCGGCG
CAGGGGCAATGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCCGGCAGGCGCGGGAGCAGGCTCTGGC
TCGGCACCTTCGACACCGCCGAGGCCGCCGCGCGCGCCAACGACGCCGCCATGATCGCGCTCTCCGCC
GGGGGCGCGGGCTGCCTCAACTTCGCCGACTCCGCCGAGCTG

>AcCBFIIIa-6

GCGAAGCGACCGGCGGGGCGGACCAAGTTCACGGAGAGGCGGCACCCGGTGTACCGCGGCGTGCGGCG
CAGGGGCAATGCCGGGCGGTGGGTGTGCGAGGTGCGCGTCCCCGGCAAGCGCGGGAGCAGGCTCTGGG
TCGGCACCTTCGACACCGCCGAGACAGCGGCGCGCGCGCACGACGCAGCCATGCTCGCCCTCGTCGCC
GGCGACGCCTGCCTCAACTTCGCAGACTCCCCCGAGCTG

>FaCBFIIIa-6

ATGAAGCGGCCGGCGGGGCGGACAAAGTTCAGGGAGACGCGGCACCCGGTGTTCCGCGGCGTGCGGCG
CAGGGGCAATGCCGGGCGTTGGGTGTGCGAGGTGCGGGTGCCAGGGCGGCGTGGAAGCAGGCTCTGGG
TCGGCACCTTCGACACCGCTGAGATCGCCGCACGCGCGCACGACGCCGCCATGCTCGCCCTGGCCGCG
GGCGATGCCTGCCTCAACTTCGCCGACTCCGCCGAGCTG

>LpCBFIIIa-6

ACGAAGCGGCCCGCGGGGCGGACAAAGTTCAGGGAGACGCGGCACCCGGTGTACCGCGGCGTGCGGCG
CAGGGGCAATGCCGGGCGGTGGGTGTGCGAGGTGCGGGTGCCAGGGCGGCGCGGGAGCAGGCTCTGGG
TCGGCACCTTCGACACTGCCGAGATCGCCGCGCGCGCACACGACGCCGCCATGCTCGCTCTCGCCGCC
GGCGACGTGTGCCTCAACTTCGCCGACTCCGCTGAGCTG

>SoCBFIIIa-6

CCGAAGCGGCCCGCGGGGCGGACCAAGTTCCGGGAGACGCGGCACCCAGTGTTCCGCGGCGTCCGGCG
CCGGGGCAACGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCCGGGAGGCGCGGCTGCAGGCTCTGGC
TCGGCACCTTCGACGCCGCCGAGGCCGCGGCCCGCGCGCACGACGCCGCCATGCTCGCCATCGCCGGA
GCGGGCCGCGCCTGCCTCAACTTCGCCGACTCGGCCTGGCTC

>SbCBFIIIa-6

CCGAAGCGGCCCGCGGGGCGGACCAAGTTCCGGGAGACGAGGCACCCGGTGTTCCGCGGCGTCCGGAG
GCGGGGCAACGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCGGGGAGGCGCGGCTGCAGGCTCTGGC
TCGGCACCTTCGACACCGCGGACGCCGCGGCCCGCGCGCACGACGCCGCCATGCTCGCCATCGCCGGC
GCGGGCGCCTGCCTCAACTTCGCCGACTCGGCCTGGCTC

>PvCBFIIIa-6

CCGAAGCGGCCCGCGGGGCGGACCAAGTTCCGGGAGACGCGGCACCCCGTGTTCCGCGGCGTGCGGCG
CCGGGGCAACGCCGGGCGGTGGGTGTGCGAGGTGCGCGTGCCGGGCCGCCGCGGCTGCAGGCTCTGGC
TCGGCACCTTCGACACCGCCGAGGGGGCCCGCGCGCGCGCACGACGCCGCCATGCTCGCCATCGCCGGC
GCGGGCGCCTGCCTCAACTTCGCCGACTCGGCGTGGCTC

>BdCBFIIIa-6

CCGAAGCGCCCCGCGGGGCGGACCAAGTTCCGGGAGACACGGCACCCGGTGTTCCGCGGCGTGCGCCG
CCGGGGCAATGCCGGCCGGTGGGTGTGCGAGGTGCGCGTGCCCGGGCGCCGCGGCTCCAGGCTCTGGC
TCGGCACCTTCGACACCGCCGAGGCCGCCGCGCGCGCGCACGACGCCGCCATGCTCGCCCTGGCCGGC
GCCGGCGCCGCCTGCCTGAACTTCGCCGACTCCGCCCAGCTG

>TmCBFIIIb-18

CCGAAGCGCCCCGCGGGCCGCACCAAGTTTCGGGAGACGCGGCACCCCGTGTTCCGCGGCGTGCGGCG
CCGCGGGAACGCCGGCCGGTGGGTCTGCGAGGTGCGCGTCCCCGGGGACCGCGGCACGCGCCTCTGGC
TCGGGACGTACTTCACCGCCGAGGCGGCCGCGCGCGCGCACGACGCCGCCATGCTCATGCTGCGTGGC
CGTTCCGCCGCGTGCCTCAACTTCCGGGACTCCGCGTGGCTG

>SoCBFIIIb

AAGAAGCGCCCCGCGGGGCGCACCAAGTTCCGTGAGACGCGGCACCCGGTGTTCCGCGGCGTGCGGCG
CCGGGGCGCCGCGGGCCGGTGGGTCTGCGAGGTGCGCGTCCCGGGGAAGCGCGGCGCGCGGCTGTGGC
TGGGGACCTACCTCGCCGCCGAGTTCGCGGCGCGCGCGCACGACGCGGCCATGCTCGCCCTGCTGGGG
CGCGGCGCCGGTGCCGGTGCGGGGCGCCTCAACTTCCCGGACTCCGCGTGGCTG

>ZmCBFIIIb-1A

AAGAAGCGCCCCGCGGGGCGCACAAAGTTCCGGGAGACGCGGCACCCGGTGTTCCGCGGCGTGCGGCG
GCGGGGCGCCGCGGGCCGGTGGGTGTGCGAGGTGCGCGTCCCGGGGAGGCGCGGCGCGCGGCTGTGGC
TCGGCACCTACCTCGCCGCCGAGGCGGCGGCGCGCGCGCACGACGCCGCGATACTCGCCCTGCAGGGC
CGCGGCGCGGGGCGCCTCAACTTCCCGGACTCCGCGCGGCTG

>ZmCBFIIIb-1B

AAGAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCGGTGTTCCGCGGCGTGCGGCG
GCGGGGCGCCGCGGGGCGGTGGGTGTGCGAGGTGCGCGTCCCCGGGCGGCGCGGCGCGCGGCTGTGGC
TCGGCACCTACCTCGGCGCCGAGGCGGCGGCGCGCGCGCACGACGCCGCGATGCTCGCCCTGGGCCGC
GGCGCGGCCTGCCTCAACTTCCCCGACTCCGCGTGGCTG

>OsCBFIIIb-1H

AAGAAGCGGCCCGCGGGGCGCACCAAGTTCAGGGAGACGCGGCACCCGGTGTACCGCGGCGTGCGGCG
CCGCGGCGGGGCGGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGGGAAGCGCGGCGCGCGCCTGTGGC
TCGGCACGTACGTCACCGCCGAGGCCGCGGCGCGCGCGCACGACGCCGCCATGATCGCGCTCCGCGGC
GGCGCCGGCGGAGGCGGCGCGGCGTGCCTCAACTTCCAGGACTCCGCGTGGCTG

>TmCBFIIIc-13

GCCAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCGGTGTACCGCGGCGTGCGGCG
CAGGGGCAACGCGGGGCGGTGGGTGTGCGAGGTGCGCGTCCCCGGCAAGCGTGGCTCTCGGCTCTGGC
TCGGGACGTACCTGACGGCCGAGGCCGCCGCCCGCGCGAACGACGCCGCCATGCTCGCACTCGGCGGC
CGCTCAGCAAGGTGCCTCAACTTCGCGGACTCGGCGTGGCTG

>HvCBFIIIc-13

GCGAAGCGCCCCGTGGGGCGCACTAAGTTCCGGGAGACCCGGCACCCGGTGTACCACGACGTGCGTCG
CAGGGGCAACGCGGGACGGTGGGTCTGCGAGGTGCGCGTCCCCAGCAAGCGTGGCGCTCGGCTCTGGC
TCGGGACGTACCTCACCGCCGGGGCCGCCGCCCGCGCCAACGACGCCGCCATGCTCGCACTCGGCGGC
CGCTCCGCCAGGCGCCTCAACTTCGCGGACTCCGCGTGGCTG

>TaCBFIIIc-3.1

GCCAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGGCACCCGGTGTACCGCGGCGTGCGGCG
CCGGGGCAACGCCGAACGTTGGGTCTGCGAGGTGCGCGTCCCCGGCAAGCGCGGCGCGAGGCTCTGGC
TCGGGACGTACGACACGGCTGAGCTCGCAGCGCGCGCGAACGACGCCGCCATGCTTGCCCTGGGCGGC
CGCTCCGCCGCGTGCCTCAACTTCGCGGACTCCGCGTGGTTG

>HvCBFIIIc-3

GCCAAGCGCCCCGCTGGGCGCACCAAGTTCCGGGAGACGCGGCACCCGGTGTACCGCGGCGTGCGGCG
CCGGGGCAACACCGAACGGTGGGTCTGCGAGGTGCGCGTCCCCGGCAAGCGCGGTGCTCGGCTCTGGC
TCGGGACGTACGCCACGGCTGAGGTCGCTGCGCGCGCGAACGACGCTGCCATGCTCGCCCTGGGCGGC
CGCTCCGCCGCGTGCCTCAACTTCGCCGACTCCGCGTGGCTG

>AstCBFIIIc-3

CGAGGCGTGCGGCGCCGGGGCAGCGCCGGCCGGTGGGTNTGCGAGGTGCGCGTCCCCGGGAGGCGTGG
CGCGCGGCTNTGGCTCGGGACGTACGCCACCGCCGAGGCCGCCGCGCGCGCGAACGACGCCGCCATGC
TCGCGCTNGGCGGCCGCTCCGCCGCGTGCCTCAACTTCGCTGACTCCGCGTGGCTG

>TaCBFIIIc-B10

GCCAAGCGCCCCGCGGGGCGCACCAAGTTCCGCGAGACGCGGCACCCTGTGTATCGTGGCGTGCGGCG
CCGGGGCAACGCCCAACGGTGGGTCTGCGAGGTGCGCGTCCCTGGCAAGCGCGGCGCTCGGCTCTGGC
TCGGGACTTACGCCACGGCCGAGATTGCAGCGCGCGCCAACGACGCCGCCATGCTCGCCCTGGGCGGC
CGCTCCGCCGCGCTCCTCAACTTCCCGGACTCCGCGTGGCTG

>TmCBFIIIc-10

GCCAAGCGCCCGGCGGGGCGCACCAAGTTCCGGGAGACAAGGCACCCGGTGTACCGCGGCGTGCGGCG
CCGGGGCAACGCCGAACGGTGGGTCTGCGAGGTGCGCGTCCCCGGCAAGCGCGGCGCGAGGCTCTGGC
TCGGGACGTATGCCACGGCCGAGATCGCAGCGCGCGCCAACGACGCCGCCATGCTCGCCCTGGGCGGC
CGCTCCGCCGCGCGGCTCAACTTCCCGGACTCCGCGTGGCTG

>HvCBFIIIc-10A

GCCAAGCGCCCGGCGGGGCGCACCAAGTTCCGCGAGACAAGGCACCCGGTGTACCGCGGCGTGCGGCG
CCGGGGCAACGCCGAACGGTGGGTATGCGAGGTGCGCGTCCCCGGCAAGCGCGGCGCTCGGCTCTGGC
TCGGGACTTACGCCACGGCCGAGATCGCAGCGCGCGCGAACGATGCCGCAATGCTCGCCCTGGGCGGC
CGCTCCGCCGCGCGCCTCAACTTCTCGGACTCCGCGTGGCTG

>HvCBFIIIc-8A

CCGAAGCGCCCCGCGGGGCGCACTAAGCTCCGGGAGACCCGGCACCTGGAGTACCGTAGCGTGCGGTG
CCGGGGCAACGCGGACGGTGGGTCTGCGAGGTGCGCGTCCCAGGCAAGCGCGGCGCTCGGCTCTCGCT
CGGGCCGTACCTCACCGCCGAGGCCGCCGCCCACGCCAACGACGCCGCCATGCTCGTGCTCGGCGGCC
GCTCCGCCAGGTGCCTCAACTTCGCGGACTCCGCGTGGCTG

>FaCBFIIIc-3

GCGAAGCGCCCCGCGGGGCGCACCAAGTTCCGGGAGACGCGCCACCCGGTGTACCGCGGCGTGAGGCG
CCGGGGCAACGCCGGCCGGTGGGTGTGCGAGGTGCGCGTCCCCGGCAAACGCGGCGCGCGGCTCTGGC
TCGGGACGTACGCCACCGCCGAGGCCGCCGCGCGCGCCAACGACGCCGCCATGCTTACGCTCGGCGGC
TGCTCCGCCTCGTGCCTCAACTTCGCCGACTCCGCCTGGCTG

>TaCBFIIId-12.1

CCGAAGCGCCCCGCGGGGCGAACCAAGTTCAAGGAGACGCGCCACCCGGTGTTCCACGGCGTGCGCCG
TCGGGGCAGCAACGGCCGGTGGGTGTGCGAGGTGCGCGTGCCGGGGAAGCGCGGCGAGCGGCTCTGGC
TCGGCACGCACGTCACCGCCGAGGCGGCCGCGCGCGCGCACGACGCCGCCATGCTCGCGCTGTACGGC
CGCACCCCCGCCGCGCGCCTCAACTACCCCGACTCGGCGTGGCTG

>HvCBFIIId-12

CCGAAGCGCCCCGCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGGTGTTCCACGGCGTGCGCCG
CCGGGGCAGCAACGGCAGGTGGGTGTGCGAGGTGCGCGTGCCGGGGAAGCGCGGCGAGCGGCTCTGGC

TCGGCACGCACGTCACCGCCGAGGCGGCCGCGCGCGCGCACGACGCCGGCATGCTCGCGCTGTACGGC
CGCACACCTGCCGCGCGCCTCAACTTCCCGGACTCGGCGTGGCTG

>TaCBFIIId-A19

GCGAAGCGGCCCGCAGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGGTGTACCGTGGCGTGCGGCG
CCGGGGCAGCGCGGGCCGGTGGGTGTGCGAGGTGCGCGTCCCCGGCAAGCGCGGCGAGCGGCTCTGGC
TCGGGACGTACGTCGCCGCCGAGTCCGCCGCGCGCGCTCACGACGCCGCCATGCTCGCCCTGCTCGGA
CGCTCCCCCTCCGCCGCGGCGTGCCTCAACTTCCCGGACTCCGCGTGGCTG

>AcCBFIIId-19

CCCGGCAAGCGCGGCGAGCGGCTCTGGCTCGGGACTACGTCGCCGCCGAGACCGCCGCGCGCGCGCAC
GACGCCGCCATGCTCGCCTTTGCGCGGCTGCTCGACCTCCGCCTCGGCGCTTGCGGACCTCAACTTCG
CGGACTC

>TaCBFIIId-A15

CTGAAGCGTCCCGCCGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGGTGTACCGCGGCGTGCGGCG
CCGCGGCAGCGCCGGCCGGTGGGTGTGCGAGGTGCGCGTCCCCGGCAAGCGCGGCGAGCGGCTGTGGC
TCGGCACGCACCTCACTGCCGAGGCGGCCGCGCGAGCGTACGACGCCGCCATGCTCTGCCTGATCGGC
CCCTCCACCCAGTGTCTCAACTTCGCCGACTCCGCCTGGCTC

>TmCBFIIId-16

CCGAAGCGCCCCGCCGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGGTGTACCGCGGCGTGCGGCG
CCGGGGCAACGCCGGCCGCTGGGTGTGCGAGGTGCGCGTCCCCGGGCAGCGCGGCGAGCGGCTCTGGC
TCGGAACGTACCTCACAGCCGACGCGGCCGCGCGTGCGCACGACGCCGCCATGCTCGGCCTGCTCGGT
CGCTCCGCCGCGTGCCTCAATTTCGCTGACTCCGCTTGGCTC

>TmCBFIIId-17

CCGAAGCGCCCCGCGGGGCGCACCAAGTTCAAGGAGACGCGCCACCCGGTGTACCGAGGCGTGCGGCG
CCGGGGCGGAGCCGGCCGCTGGGTCTGCGAGGTGCGCGTGCCCGGCAGGCGCGGGTGCAGGCTCTGGC
TCGGGACCTACGTCACCGCCGAGTCCGCCGCGCGCGCGCACGACGCCGCCATGCTCGCGCTGGGCGGC
CGCTCCGCCGCGTGCCTCAACTTCCCGGACTCCGCGTGGCTG

>AsCBFIIId-16A

CCGAAGCGCCCAGCGGGGCGCACCAAGTTCAAGGAGACGCGGCACCCGGTGTACCGCGGCGTGCGACG
CCGCGGCAACGCCGGGCGATGGGTGTGCGAGGTGCGCGTCCCCGGACAGCGCGGGGAGCGGCTCTGGC
TCGGGACGTACCTCACCGCCGAGTCCGCCGCGCGCGCGCACGACGCCGCCATGCTCGGGCTCCACGGC
CGCTCCGCCTCCGCATGCCTCAACTTCGCCGACTCCGCGTGGCTT

>AsCBFIIId-16B

CCGAAGCGCCCAGCGGGGCGCACCAAGTTCAAGGAGACGCGGCACCCGGTGTACCGCGGCGTGCGGCG
CCGCGGCAACGCCGGCCGGTGGGTGTGCGAGGTGCGCGTCCCAGGGCAGCGCGGCGAGCGGCTGTGGC
TCGGGACGTACCTCACCGCCGAGTCCGCCGCGCGTGCGCACGACGCCGCCATGCTCGGGCTGCTCGGC
CGCTCCGCATCCTCCGCGTGCCTCAACTTCGCCGACTCCGCGTGGCGC

>AsCBFIIId-12

CCGAAGCGCCCCGCGGGACGCACCAAGTTCAAGGAGACGCGCCACCCTGTGTACCGCGGCGTGCGTCG
CCGCGGCAGCAACGGGCGGTGGGTGTGCGAGGTGCGCGTGCCCAGCAAGAGCGGCGAGCGTCTCTGGC
TCGGCACGCACGTCACGGCCGAGGCCGCCGCGCGCGCGCACGACGCCGCCATGCTCGCGATGCACGGC
CACACTTCCGCCGCGTGCCTCAACTTCCCGGACTCCGCGTGGCTG

>OsCBFIII-1D

CCGAAGCGGCCGGCGGGGCGAACCAAGTTCCAGGAGACGAGGCACCTAGTGTTCCGTGGGGTGCGATG
GCGTGGGTGCGCGGGGCGGTGGGTGTGCAAGGTGCGTGTCCCGGGCAGCCGCGGTGACCGTTTCTGGA
TAGGCACGTCTGACACCGCCGAGGAGACCGCGCGCACGCACGACGCCGCCATGCTCGCCTTGTGCGGG
GCCTCCGCCAGCCTCAACTTCGCCGACTCTGCCTGGCTG

>OsCBFIII-1I

AGCAAGCGCCCCGCGGGGCGCACCAAGTTCCACGAGACCCGCCACCCGGTGTTCCGCGGCGTGCGGCG
CCGCGGCCGCGCGGGGCGGTGGGTGTGCGAGGTCCGCGTGCCGGGCCGCCGCGGGTGCAGGCTCTGGC
TCGGCACGTTCGACGCCGCCGACGCCGCCGCGCGCGCCCACGACGCCGCCATGCTCGCGCTCCGCGGC
CGCGCCGCCGCGTGCCTCAACTTCGCCGACTCCGCCTGGCTG

>OsCBFIII-1J

CCGAAGCGGCCGGCGGGGCGGACCAAGTTCCAGGAGACGAGGCACCCGGTGTTCCGCGGGGTGCGGCG
GCGCGGGCGCGCGGGGCGGTGGGTGTGCGAGGTGCGCGTCCCGGGCAGCCGCGGCGACCGCCTGTGGG
TCGGCACGTTCGACACCGCCGAGGAGGCCGCGCGCGCGCACGACGCCGCCATGCTCGCCCTGTGCGGG
GCCTCCGCCAGCCTCAACTTCGCCGACTCCGCCTGGCTG

>SoCBFIV

CCCAAGAAGCCGGCGGGACGGACCAAGTTCCGGGAGACGCGGCATCCAGTGTACCGCGGCGTCCGGCG
CCGCGGCACACTGGGGCGGTGGGTGTGCGAGGTGCGCATCCCCGGGGGGCGCGGGTCAAAGCTCTGGC
TCGGCACCTTCGCCACGCCCGACCTGGCGGCGCGCGCGCACGACGCCGCCGACATCGCGCTCTCGGGC
CGCGGCGCCTGCCTCAACTTCGCCGACTCAGCGTGGCTG

>OsCBFIV-1B.1

CCGAAGAGGCCGGCGGGAAGGACCAAGTTCAGGGAGACGAGGCACCCGGTGTACCGCGGCGTGCGGCG
GCGCGGGGGGCGGCCGGGCGCGGCGGGGAGGTGGGTGTGCGAGGTGCGGGTGCCCGGGGCGCGCGGCT
CCAGGCTGTGGCTCGGCACGTTCGCCACCGCCGAGGCGGCGGCGCGCGCGCACGACGCCGCCGCGCTG
GCGCTCCGCGGCAGGGCCGCCTGCCTCAACTTCGCCGACTCCGCGTGGCGG

>FaCBFIVa-2

CCGAAGCCACGATCGGGGCGCACCAAGTTCCAGGAGACGCGGCACCCGGTGTACCGCGGCGTTCGGCG
CCGGGGGCGTGCCGGGCAGTGGGTGTGCGAGATGCGCGTCCACGGGACGAAGGGGTCCAGGCTCTGGC
TCGGCACCTTCGACACCGCTGAGATGGCTGCACGCGCGCACGACGCCGCCGCGCTCGCGCTCTCCGGC
CGCGACGCATGCCTCAACTTCGCTGACTCCGCCTGGCGG

>TaCBFIVa-A2

CCGAAGCGGCGGGCGGGGCGGAACAAGTTGCAGGAGACACGCCACCCAGTGTACCGCGGCGTGCGCCG
CCGTGGCCGGGAAGGGCAGTGGGTGTGCGAGCTGCGCGTGCCGGCCGGAAGCCGGAGTTACTCCAGGA

TCTGGCTTGGCACCTTCGCCAGTGCCCAGATGGCGGCGCGCGCGCACGACTCGGCCGCGCTCGCGCTC
TCCGGCCGCGACGCGTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>TmCBFIVa-2

CCGAAGCGGCGGGCGGGGCGGAACAAGTTGCAGGAGACGCGCCACCCGGTGTACCGCGGCGTGCGCCG
CCGTGGCCGGGAAGGGCAGTGGGTGTGGGTGTGCGAGCTGCGCGTGCCGGCTGCCGGAAGCCGGGTTT
ACTCCAGGATCTGGCTCGGCACCTTCGCCGACCCCGAGATGGCGGCGCGCGCGCACGACTCGGCCGCG
CTCGCGCTCTCCGGCCGGGATGCGTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>HvCBFIVa-2A

CCGAAGCGGCGGGCCGGCCGGATCAAGTTGCAGGAGACGCGCCACCCGGTGTACCGCGGCGTGCGACG
CCGTGGCAAGGTCGGGCAGTGGGTGTGCGAGCTGCGCGTCCCCGTAAGCCGGGGTTACTCCAGGCTCT
GGCTCGGCACCTTCGCCAACCCCGAGATGGCGGCGCGCGCGCACGACTCCGCCGCGCTCGCCCTCTCC
GGCCATGATGCGTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>HbCBFIVa-2

CCGAAGCGGCGGGCCGGCCGGATCAAGTTGCAGGAGACGCGCCACCCGGTGTACCGCGGCGTGCGACG
CCGTGGCAAGGTCGGGCAGTGGGTGTGCGAGCTGCGCGTCCCCGTAAGCCGGGGTTACTCCAGGCTCT
GGCTCGGCACCTTCGCCAACCCTGAGATGGCGGCGCGCGCGCACGACTCCGCCGCGCTCGCGCTCTCC
GGCCACGATGCTTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>AsCBFIV

TGGGTGTGCGAGATGCGCGTCCAGGGATGCAGGGGCTCTAGGCTCTGGCTCGGCACCTTCGCCTCAGC
CGAGATGGCGGCCCGCGCGCACGACGCCGCCGCGCTCGCGCTCTCCGGGCCCGACGCCTGCCTCAACT
TTGCCGACTCCGCATGGCGG

>ScCBFIVb-20

CCCAAGAGGCCGGCAGGGCGGACCAAGTTTAAGGAGACCCGCCACCCGCTGTACCGCGGCGTGCGGCG
GCGGGGTCGGCTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCGCGGCGCGCAAGGGTACAGGCTCTGGC
TCGGCACATTCACCACCGCCGAGATGGCGGCGCGCGCGCACGACTCCGCCGTGCTCGCGCTCCTCGAC
CGCGCCGCTTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>TaCBFIVb-A20

CCCAAGAGGCCGGCAGGGCGGACCAAGTTCAAAGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCG
GCGGGGCCGGCTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCGCGGCGCGCAAGGGTACAGGCTCTGGC
TCGGCACCTTCACCACGGCCGAGATGGCGGCGCGCGCGCACGACTCCGCCGTCCTCGCGCTCCTCGAC
CGCGCCGCCTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>TaCBFIVb-21.1

CCGAAGAGGCCGGCGGGGCGGATCAAGTACAAGGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCG
TCGGGGGCGGTACGGGCGGTGGGTGTGCGAGGTGCGCGTGCGCGGCACCAAGGAGACAAGGCTCTGGC
TCGGCACCTTCCGCACCGCCGAGATGGCGGCGCGAGCGCACGACTCCGCCTCGCTCGCTCTCTCCGGA
AGCGCCGCATGCCTCAACTTCGCCGACTCCGCCTGGCGG

>HvCBFIVc-14

CCGAAGCGGCCTGCCGGCCGGACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCG
CCGGGGCCCCGCAGGCCGGTGGGTGTGTGAGGTGCGCGTGCTCGGGATGAGGGGCTCCAGGCTCTGGC
TCGGCACCTTCACCACCGCCGAGATGGCAGCGCGCGCGCACGACGCCGCCGTTCTCGCGCTCTCTGGC
CGCGCCGCCTGTCTGAACTTCGCCGACTCTGCTTGGCGG

>TaCBFIVc-14.1

CCGAAGCGGCCGGCCGGCCGGACCAAGTTCAAGGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCG
ACGGGGCCCCGCCGGCCGGTGGGTGTGCGAGGTGCGCGTGCTCGGGATGAGGGGCTCCAGGCTTTGGC
TCGGCACCTTCACCACCGCCGAGATGGCAGCGCGCGCGCACGACGCCGCCGTTCTCGCGCTCTCCGGC
CGCGCCGCTTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>HvCBFIVd-4A

CCGAAGCGGCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCG
CCGGGGCCGGGTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGC
TCGGGACCTTCACCAACCCCGAGATGGCCGCGCGCGCCCACGACGCCGCGGTGCTCGCGCTCTCCGGC
CGCGCCGCCTGCCTCAACTTCGCCGACTCCGCGTGGCGG

>TmCBFIVd-4

CCGAAGCGGCCCGCGGGGCGGACCAAGTTCCAGGAGACGCGCCACCCGCTGTACCGCGGCGTGCGCCA
GCGGGGGCCCGCCGGGCGGTGGGTCTGCGAGGTGCGCGTGCTCGGGATGAGGGGATCCAGGCTCTGGC
TCGGCACCTTCGTCACCGCCGAGATGGCGGCGCGCGCCCACGACGCCGCCGTGCTCGCGCTCTCCGGC
CGCAAAGCCTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>TaCBFIVd-4.1

CCGAAGCGGCCGGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGGGGCGTGCGGCG
CCGTGGCCGGGTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGGTGAAGGGCTCCAGGCTCTGGC
TCGGCACCTTCACCACCGCCGAGATGGCGGCGCGCGCGCACGACGCCGCGGTGCTCGCGCTCTCCGGC
CGCGCCGCCTGCCTCAACTTCGCCGACTCCGCATGGCGG

>TaCBFIVd-9.1

CCGAAGCGGCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCG
CCGTGGCCGCGTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGC
TCGGCACCTTCAACACGGCCGAGATGGCGGCGCGCGCGCACGACGCCGCCGTGCTCGCGCTCTCCGGC
CGCGCCGCCTGCCTCAACTTCGCCGACTCCGCCTGGCGC

>ScCBFIVd-9A

CCGAAGCGCCCCGCGGGGAGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCG
CCGTGGCCGCGTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGC
TCGGCACCTTCAACACGGCCGAGATGGCGGCGCGCGCGCACGACGCTGCCGTGCTCGCGCTCTCCGGC
CGCGCCGCCTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>ScCBFIVd-9B

CCGAAGCGCCCCGCGGGGCGGACCAAGTTCCACGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCG
CCGTGGCCGCGTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGC
TCGGCACCTTCAACACGGCCGAGATGGCGGCGCGCGCCCACGACGCAGCCGTGCTCGCGCTCTCCTGC
CGCGCCGCCTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>HvCBFIVd-9

CCGAAGCGGCCCGCGGGGCGCACCAAGTTCCACGAGACGCGCCACCCTCTGTACCGCGGCGTGCGCCG
CCGCGGCCGCGTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGATCAAGGGCTCCAGGCTCTGGC
TCGGCACCTTCAACACGGCCGAGATGGCCGCCCGCGCGCACGACGCCGCCGCGCTCGCGCTCTCCGGC
CGCGCCGCCTGCCTCAACTTCGCCGACTCCGCCTGGCGG

>TaCBFIVd-A22

CCGAAGCGCCCCGCGGGGCGGACCAAGGTGCATGAGACGCGCCACCCGCTGTACCGCGGCGTGCGGCA
GCGGGGCCGGGTCGGGCAGTGGGTGTGCGAGGTGCGCGTGCCCGGGGTGAAGGGCTCCAGGCTCTGGC
TCGGCACCTTCGCCACCGCCGAGATGGCGGCGCGCGCGCACGACGCAGCCGTGCTCGCGCTCTCCGGC
CGCGCCGCCTGCCTCAACTTCGCCGACTCCGCGTGGCGC

III-Table S6: List of protein sequences of monocot CBFs used for the hydrophobic cluster analysis

>AtCBF3

MNSFSAFSEMFGSDYESSVSSGGDYIPTLASSCPKKPAGRKKFRETRHPIYRGVRRRNSGKWVCEVRE
PNKKTRIWLGTFQTAEMAARAHDVAALALRGRSACLNFADSAWRLRIPESTCAKDIQKAAAEAALAFQ
DEMCDATTDHGFDMEETLVEAIYTAEQSENAFYMHDEAMFEMPSLLANMAEGMLLPLPSVQWNHNHEV
DGDDDDVSLWSY

>OsCBFI-1F

MDTEDTSSASSSSVSPPSSPGGGHHHRLPPKRRAGRKKFRETRHPVYRGVRARAGGSRWVCEVREPQA
QARIWLGTYPTPEMAARAHDVAAIALRGERGAELNFPDSPSTLPRARTASPEDIRLAAAQAAELYRRP
PPPLALPEDPQEGTSGGGATATSGRPAAVFVDEDAIFDMPGLIDDMARGMMLTPPAIGRSLDDWAAID
DDDDHYHMDYKLWMD

>ZoCBFI-1

MDSDSLDSPTLPRRSASDEEVCYATVSSAPPKRRAGRTKFRETRHPVFKGVRRRNGDKWVCEVREPNK
KSRIWLGTFPTAEMAARAHDVAAIALRGRAACLNFADSAWLLPSPDSSSPIDIRKAAALAAEAFRPNS
TPSTALAPAAAAADPSTSWAHAGPSSTPSSDPLFDDGDFNFGMQDYLDMAEGLLIGPPPPPPAPANDE
DYDDVSLWSYSI

>RhCBFI-1

MENFSDYSMDSPLAQRSASDEEAYATVSSAPPKRRAGRTKFRETRHPVYKGVRRRNADKWVCEVREPN
KKSRIWLGTFPTAEMAARAHDVAAMALRGRSACLNFADSAWLCPVPGSSNPKDIQRAAVDAAEAFRPQ
TERVDAAESREDAAMAMIARPSLAAADGPFYMEDGLNFGMQGYLDMAAGMLMEPPPIYADEEVSDGDV
SLWSYSI

>HvCBFIa-1

MDVGALSSDYSSGTPSPVGADGGNSEGFSTYMTVSSAPPKRRAGRTKFKETRHPVYKGVRRRNPGRWV
CEVREPHSKQRIWLGTFETAEMAARAHDVAALALRGRAACLNFADSPRRLRVPAVGASPDEIRRAAVE
AAEAFLPAPDQSNAPAEEVAAAPTMQFAGDPYYGMDDGMDFGMQGYLDMAQGMLIAPPPLVGPSATAG
DGDDDGEVSLWSY

>OsCBFIa-1G

MDVSAALSSDYSSGTPSPVAADADDGSSAYMTVSSAPPKRRAGRTKFKETRHPVFKGVRRRNPGRWVC
EVREPHGKQRIWLGTFETAEMAARAHDVAALALRGRAACLNFADSPRRLRVPPIGASHDDIRRAAAEA
AEAFRPPPDESNAATEVAAAASGATNSNAEQFASHPYYEVMDDGLDLGMQGYLDMAQGMLIDPPPMAG
DPAVGSGEDDNDGEVQLWSY

>TaCBFIa-A11

MEWAYSGGGHSSSGTKSPAAGGREEGSYMTVSSAPPKRRAGRTKVRETRHPVYKGVRSRNPGRWVCEV
REPHGKQRLWLGTFDTAEMAARAHDVAALALRGRAACLNFADSPRTLRVPPQGAGHEEIRRAAVEAAE
LFRPEPGQRNAATTEAPAASPADAGNAELVANSPYHLMDGLEFEMQGYLDMAHGMLIEPPPMAGPSTW
IEEDYDCEVSLWNY

>OsCBFIa-1E

MEWAYYGSGYSSSGTPSPVGGDGDEDSYMTVSSAPPKRRAGRTKFKETRHPVYKGVRSRNPGRWVCEV
REPHGKQRIWLGTFETAEMAARAHDVAAMALRGRAACLNFADSPRRLRVPPLGAGHEEIRRAAVEAAE
LFRPAPGQHNAAAEAAAAVAAQATAASAELFADFPCYPMDGLEFEMQGYLDMAQGMLIEPPPLAGQST
WAEEDYDCEVNLWSY

>TaCBFII-5.1

MDQYSYRGGGDDNGQGGYATVTSAPPKRPAGRTKFRETRHPVYRGVRRRGAAGRWVCEVREPNKKSRI
WLGTFASPEAAARAHDVAALALRGRAACLNFADSAALLAVDPATLRTPDDIRAAAIALAETACPAAPA
SSSAVAAVASAPAPPMMTMMHESAAVHYDDYPMQYGYGGIGDLDQDSYYYDGMSAAGGDWQSGSHMDG
ADDDCNDSGGYGAGEVPLWSY

>TmCBFII-A5

MDNSGVVFYGGAYATVMSAPPKRPAGRTKFRETRHPVYRGVRRRGAAGRWVCEVRQPNNKSRIWLGTF
ASPEAAARAHDVAALALRGRAACLNFADSAALLAVDPATLRTPQDIRAAAITLAQTACPHDAPRSSVS
AASAPAPAMVITQEAAAAPYDSYAMYGGLADLEQHSHCYYDGMSGSGDWQSISHMNVADEDGGYGAGD
VALWSY

>OsCBFII-1C

MEYYEQEEYATVTSAPPKRPAGRTKFRETRHPVYRGVRRRGPAGRWVCEVREPNKKSRIWLGTFATAE
AAARAHDVAALALRGRGACLNFADSARLLRVDPATLATPDDIRRAAIELAESCPHDAAAAAASSSAAA
VEASAAAAPAMMMQYQDDMAATPSSYDYAYYGNMDFDQPSYYYDGMGGGGEYQSWQMDGDDDGGAGGY
GGGDVTLWSY

>TaCBFIIIa-6.1

MCPIKKEMSGESGSPCSGENFYSPSTSREHQQAKQAAWTSAPAKRPAGRTKFRETRHPVYRGVRRRGN
AGRWVCEVRVPGRRGSRLWLGTFDTAEAAARANDAAMIALSAGGAGCLNFADSAELLAVPAASSYRSL
DEVRHAVVEAVEDFLRREAIAEEDALSGTSSSAPSSLTDDESSSSPPEDSPFELDVLSDMGWDLYYAS
LAQAMLMAPPSSMAAALGDYGEVDVPLWSYQS

>TaCBFIIIa-6.2

MCPIKREMSGESGSPSPCSGENFCSPSASPERQQARQAGWTSAPAKRPAGRTKFRETRHPVYRGVRRR
GNAGRWVCEVRVPGRRGSRLWLGTFDTAEAAARANDAVMLMLAAGGAACLNFADSAELLSVPVASSYR
SLDEVRHAVVEAVEDLLRREALAEEDALSGTSSSAPSPLTDDESSSSPLPEEDSPFEQDVLSEMGWDL
YYASLAQAMLMAPPAAAAALGDYGEAHLADVPLWSYQS

>OsCBF3A-1A

MCGIKQEMSGESSGSPCSSASAERQHQTVWTAPPKRPAGRTKFRETRHPVFRGVRRRGNAGRWVCEVR
VPGRRGCRLWLGTFDTAEGAARAHDAAMLAINAGGGGGGGACCLNFADSAWLLAVPRSYRTLADVRHA
VAEAVEDFFRRRLADDALSATSSSSTTPSTPRTDDDEESAATDGDESSSPASDLAFELDVLSDMGWDL
YYASLAQGMLMEPPSAALGDDGDAILADVPLWSY

>TmCBFIIIb-A18

MDMSLEHSSSASSSSTTERGGTAWPWPPKRPAGRTKFRETRHPVFRGVRRRGNAGRWVCEVRVPGDRG
TRLWLGTYFTAEAAARAHDAAMLMLRGRSAACLNFRDSAWLLSVPPAFSNLSDVRRAAVQAVADFLRR
PEATGAFAGAAQEVTSSVTVPSAAACSVPSSETAQTSGDANFEEPGALSMDMFDLDCLFGETDSDTYY
YANLAQGLLMEPPPSMATGAYWDNGDCADGGAGADVALWSY

>OsCBFIIIb-1H

MDMAGHEVNSSSSSSGAESSSSSSGRQQYKKRPAGRTKFRETRHPVYRGVRRRGGAGRWVCEVRVPGK
RGARLWLGTYVTAEAAARAHDAAMIALRGGAGGGGAACLNFQDSAWLLAVPPAAPSDLAGVRRAATEA
VAGFLQRNKTTNGASVAEAMDEATSGVSAPPPLANNAGSSETPGPSSIDGTADTAAGAALDMFELDFF
GEMDYDTYYASLAEGLLMEPPPAATALWDNGDEGADIALWSY

>TaCBFIIIc-D3

MDMGLEVSSSSPSSSSVSSSPVHAAGRASLAKRPAGRTKFRETRHPVYRGVRRRGNAERWVCEVRVPG
KRGARLWLGTYATAEVAARANDAAMLALGGRSAACLNFADSAWLLAVPPALSDLGDVRRAAVEAVADF
QRREAANGSLTATVTEEASCGAPEESSSESDSAGSSETSEPSADAEFEVPVAVDTDMFSRLDLFPEMD
LCSYYASLAEALLVDPPSTVAIIDSYWDNGDDGADVALWSY

>TaCBFIIIc-B10

MDMGLEVSSSSPSSSSLAKRPAGRTKFRETRHPVYRGVRRRGNAQRWVCEVRVPGKRGARLWLGTYAT
AEIAARANDAAMLALGGRSAALLNFPDSAWLLAVPSAHSDLADVRRAAVEAVADLQRREAAGGSITAT
ATATAAEEASCGAPAESSSESDDAGSSETSKPSADGDFAVPGGMDIEMFSRLDLFPEMDLGSYYASLA
EALLMDPPPVATGTGAYWDNGECGEAEGATEFALWS

>TmCBFIIIc-A13

MDLSSSSPSSSASSSPEHASGRASPAKRPAGRTKFRETRHPVYRGVRRRGNAGRWVCEVRVPGKRGSR
LWLGTYLTAEAAARANDAAMLALGGRSARCLNFADSAWLLAVPSALSDLADVRRAALQAVADFQRWEA
ANGLVTRTAAEQAPSSAPAQSSSESDSADSSETSEASADGEFEVLATMDIDMFRLDLFPEMDLGSYYV
SLAEALLMDPPSTATIIDAYRDNRDGGADVALWSY

>TaCBFIIId-12.1

MDTGPERNWNSPASPPSSLEQGMPTSPASPTPKRPAGRTKFKETRHPVFHGVRRRGSNGRWVCEVRVP
GKRGERLWLGTHVTAEAAARAHDAAMLALYGRTPAARLNYPDSAWLLAVPSSLSDLADVRRAAIGAVV
DFLRRQEAGASAGAVAEEAHVDGIASAASAPDNASSSAAAAHSQPPCANAGYEVPDALCHDMFELHTS
GEMDAGTYYADLAQGLLLEPPPPPSSGASSERGDDAALWNH

>TaCBFIIId-A15

MDMTGSDQQRSSPSSPSSSSHLKRPAGRTKFKETRHPVYRGVRRRGSAGRWVCEVRVPGKRGERLWLG
THLTAEAAARAYDAAMLCLIGPSTQCLNFADSAWLLAVPSALPDFADVRRAALSAVADFQRREAASGA
ATRSLDATVPVDDGTCSQSAQSSMENTGSSWTSSSLPSGNGMFEVPATLGCDMFELDMSGEMDLDTYY
AYFAEGLLLEPPQPPVAGACWDTEGGGADAALWSY

>TmCBFIIId-A16

MPLVQTASGKTIKQCTPQDTKILTLPSQAQPALTLHRPPSTVRSSSSQHRPPSAMDMTGSDQQWSSSS
SPSSTSSHPKRPAGRTKFKETRHPVYRGVRRRGNAGRWVCEVRVPGQRGERLWLGTYLTADAAARAHD
AAMLGLLGRSAACLNFADSAWLLAVPPALADLAAVRRAALAAVADFQRRHASNSAATVPADEETSGAS

ALSSADNASGSSATSQPWAEGTFEVPSALGSDMFELDLSGEMDLGTYYADLADGLLLEPPPSLDSGAC
WDTGDGGADSGLWSY

>TmCBFIIId-A17

MDMGSEQWSSPSTSASSRDQHAAAPPKRPAGRTKFKETRHPVYRGVRRRGGAGRWVCEVRVPGRRGCR
LWLGTYVTAESAARAHDAAMLALGGRSAACLNFPDSAWLLAVPCALADLADVRRAALAAVAGFQRREA
ASGAATVPVDEVFDTSSADDAGSWSWATPQPSCAAADGMFEVPAAALASDMFDFEFDVSWVMDLGSPA
TSQPGCADKVLEVPAAALGGGDMFEFDLELDMSGEMNLVGSYYADFAEGLLLEPPQPADATEARWRNG
DYCGGDGGGDAALWSQ

>*Ta*CBFIIId-A19

MDFGINGWISSPSSSTSGHELGDAVPVWSPAAKRPAGRTKFKETRHPVYRGVRRRGSAGRWVCEVRVP
GKRGERLWLGTYVAAESAARAHDAAMLALLGRSPSAAACLNFPDSAWLLVMPPRLSDLADVRRAAIQA
VAGFLRPEAATVVPDVDEATSPVYLPSPVDNADEVFQVPTFSPLGSDMFELDMSGEMDLDAYYAGFAQ
GMLLEPPPTPAYWETGECGDGGAAAGLWSY

>OsCBFIII-1D

MEKNTAASGQLMTSSAEATPSSPKRPAGRTKFQETRHLVFRGVRWRGCAGRWVCKVRVPGSRGDRFWI
GTSDTAEETARTHDAAMLALCGASASLNFADSAWLLHVPRAPVVSGLRPPAARCATRCLQGHRRVPAP
GRGSTATATATSGDAASTAPPSAPVLSA

>OsCBFIII-1I

MCTSKLEEITGEWPPPALQAASTTSSSEPCRRLSPPSSKRPAGRTKFHETRHPVFRGVRRRGRAGRWV
CEVRVPGRRGCRLWLGTFDAADAAARAHDAAMLALRGRAAACLNFADSAWLLAVPPPATLRCAADVQR
AVARALEDFEQRESSSSVFPLAIDVVAEDAMSATSEPSAASDDDAVTSSSSTTDADEEASPFELDVVS
DMGWSLYYASLAEGLLMEPPASGASSDDDDDAIVDSSDIADVSLWSY

>OsCBFIII-1J

MEKNTTAMGQLMSSSATTAATATGPASPKRPAGRTKFQETRHPVFRGVRRRGRAGRWVCEVRVPGSRG
DRLWVGTFDTAEEAARAHDAAMLALCGASASLNFADSAWLLHVPRAPVASGHDQLPDVQRAASEAVAE
FQRRGSTAATATATSGDAASTAPPSSSPVLSPNDDNASSASTPAVAAALDHGDMFGGMRTDLYFASLA
QGLLIEPPPPPTTAEGFCDDEGCGGAEMELWS

>TaCBFIVa-A2

MDTNAAWPQFDGQEYRTVWPEEQEYRTVWSEPPKRRAGRNKLQETRHPVYRGVRRRGREGQWVCELRV
PAGSRSYSRIWLGTFASAQMAARAHDSAALALSGRDACLNFADSAWRMMPVHAAGSFKLAAAQEIKDA
VAVALKEFQEQQRPADESTAPSSTAEESALSIIPSDLSGLDNEHWIGGMEAGSYYASLAQGMLMEPPA
DGAWREDREHDDGFDTSLWSY

>TmCBFIVa-A2

MDTAGAWPHFEGQEYRTVWPEEEYRTVWSEPPKRRAGRNKLQETRHPVYRGVRRRGREGQWVWVCELR
VPAAGSRVYSRIWLGTFADPEMAARAHDSAALALSGRDACLNFADSAWRMMPVHAAGSFKLAAAQEIK
DAVAVALKAFQEQQRPADASKAPSSTDSTSEESAPSITSNDLSGLDDEHWIGGMDAGSYYANLAQGML
MEPPAAGAWREDREQDDGVDTSLWSYWLDGFGCVKL

>HvCBFIVa-2A

MDTVAAWPQFEEQDYMTVWPEEQEYRTVWSEPPKRRAGRIKLQETRHPVYRGVRRRGKVGQWVCELRV
PVSRGYSRLWLGTFANPEMAARAHDSAALALSGHDACLNFADSAWRMMPVHATGSFRLAPAQEIKDAV
AVALEVFQGQHPADACTAEESTTPITSSDLSGLDDEHWIGGMDAGSYYASLAQGMLMEPPAAGGWRED
DGEHDDGFNTSASLWSY

>TaCBFIVb-A20

MDTAAPGSPREGHRTVCSEPPKRPAGRTKFKETRHPLYRGVRRRGRLGQWVCEVRVRGAQGYRLWLGT
FTTAEMAARAHDSAVLALLDRAACLNFADSAWRMLPVLAAGSSRFSSAREIKDAVAVAVMEFQRQRPV
LSTPETHDGEKDVQGSPTPSELSTSSDLLDEHWFGGMNAGSYYASLAQGMLMEPPAARARSEDGGEYS
GVQTPLWNTYPTN

>TaCBFIVb-21.1

MDADAASPSDQHRTVWTEPPKRPAGRIKYKETRHPLYRGVRRRGRYGRWVCEVRVRGTKETRLWLGTF
RTAEMAARAHDSASLALSGSAACLNFADSAWRMLPVLAAGSSSFSSAREIKDAVAVAVVAFQRQRSVA
STADGEKDVQGSPTPSELSTSSDLLDEHWFGGTDAGSYYSPGMFMEPPERPGNRELGAGEVETLLW

>TaCBFIVc-14.1

MDAADAASPCDGHRTVWSEPPKRPAGRTKFKETRHPLYRGVRRRGPAGRWVCEVRVLGMRGSRLWLGT
FTTAEMAARAHDAAVLALSGRAACLNFADSAWRMLPVLAGPFSTAKEIKDAVAVAVLAFQRQHPVASM
APLSPARTTDDEKEIDGLPAPSALSMSSELLNEHWFGGMDAGSCYSEFMESPDTRPWREDFELGGVET
PPWSYLFD

>HvCBFIVc-14

MDAADAASPCDGHRTVWSEPPKRPAGRTKFKETRHPLYRGVRRRGPAGRWVCEVRVLGMRGSRLWLGT
FTTAEMAARAHDAAVLALSGRAACLNFADSAWRMLPLLAGPFSTAKEIKDAVAVAVLAFQRQHPVAST
APMSPARTAVDEKEVDGSPAPSALFMSSELLNEHWFGGMDAGSCYSEGMFIESPDTRPWREDLELGGV
QTPPWSYLFD

>TaCBFIVd-4.1

MDVADAASKSGQHEQGHRTVSSEPPKRPAGRTKFHETRHPLYRGVRRRGRVGQWVCEVRVPGVKGSRL
WLGTFTTAEMAARAHDAAVLALSGRAACLNFADSAWRMLPVLAAGSFGFGSAREIKLAVAVAVVAFQQ
QQIILPVACPTVEAAASPSNSLFYMSSVDLLELDEEQWFGGMDAGSYYESLAQGMLMAPPDDRARRED
AEQTGVETPTPLWSYLFD

>TmCBFIVd-A4

MPSGQEEQRHRTVRSEPPKRPAGRTKFQETRHPLYRGVRQRGPAGRWVCEVRVLGMRGSRLWLGTFVT
AEMAARAHDAAVLALSGRKACLNFADSAWRMLPVLAAGSFGFGSAREIKTAVAVAVLAFQRQQIVLPV
ACPAAEPAVAPSGALFSMSSGDLLELDDEQWFGGMVAGSYYESLAQGMLVEPPDAGAWREDSEHSGVA
ETQTPLWS

>TaCBFIVd-9.1

MDVADIASPSGQQEQGHRTVSSEPPKRPAGRTKFHETRHPLYRGVRRRGRVGQWVCEVRVPGIKGSRL
WLGTFNTAEMAARAHDAAVLALSGRAACLNFADSAWRMLPVLAAGSFGFGSASEIKAAVAVAVVAFQR
KQIVLPVAVAVVALQQKQVPIAVAVVALQQKQVPVAVAVVALQQLPVPVAVAVVALQQQQIILPVACL
APEFYMSSGDLLELDEEQWFGGMEAGSYYASLAQGMLVAPPDERARPESGEQSGVQTPLWSCLFD

>TaCBFIVd-A22

MDVADAASSSGQEQQGHRTVSSEPPKRPAGRTKVHETRHPLYRGVRQRGRVGQWVCEVRVPGVKGSRL
WLGTFATAEMAARAHDAAVLALSGRAACLNFADSAWRMLPVLAAGSFGFGSAREIKAAVAVAVVAFQK
EQIIPVAVAVVALQKQQIIPVAVAVVALQKQQIPVAVALVALQEQQVPVAVAVVALHRQQVPVACPAT
SGPGSALFYMSSSDLLELDEEQWFGGMEAGSYYASLAQGMLVAPPDERARPEDGEQSGVQTPLWSQSH
LFN

>OsCBFIV-1B.1

MEVEEAAYRTVWSEPPKRPAGRTKFRETRHPVYRGVRRRGGRPGAAGRWVCEVRVPGARGSRLWLGTF
ATAEAAARAHDAAALALRGRAACLNFADSAWRMPPVPASAALAGARGVRDAVAVAVEAFQRQSAAPSS
PAETFANDGDEEEDNKDVLPVAAAEVFDAGAFELDDGFRFGGMDAGSYYASLAQGLLVEPPAAGAWWE
DGELAGSDMPLWSY

ARTICLE III

STRUCTURE AND FUNCTIONAL ANALYSIS OF WHEAT ICE (INDUCER OF CBF EXPRESSION) GENES

Mohamed Badawi[1], Yedulla Venkat Reddy[1], Zahra Agharbaoui, Yoko Tominaga, Mario Houde and Fathey Sarhan*

Département des Sciences Biologiques, Université du Québec à Montréal, Case Postale 8888, Succursale Centre-ville, Montréal, Québec, Canada H3C 3P8

[1] These authors contributed equally to this work.

(Article in preparation for publication)

* To whom correspondence should be addressed.

Contribution

I took part in all the stages of experimentation and drafting of the article under the direction of FS. I carried out the identification, the cloning and the sequence analyses of the promoter of *TaCBFIVd-B9*, as well as the experiments of freezing tolerance. Also, I did the molecular characterization of transgenic plants. YR carried out the identification of the two ICE genes studied and gel shift assays. ZA directed and carried out the transitient expression assays and production of the transgenic plants and I assisted her in the cloning, phenotype and molecular characterization. I participated in the phylogenetic analysis with YT. All the authors took an active part in drafting the article.

Résumé

Deux gènes inducteurs d'expression des *CBF* (les gènes *ICE1*-like), *TaICE87* et *TaICE41*, ont été isolés chez le blé. Les deux gènes sont constitutivement exprimés chez le blé du printemps et le blé d'hiver, et les protéines qu'ellent encodent possèdent 94% d'homologie au niveau de leur domaine bHLH. Des analyses de retard sur gel (EMSA) démontrent que TaICE87 et TaICE41 se lient différentiellent aux élements MYC du promoteur *TaCBFIVd-B9* chez le blé. Les essais d'expression transitoire chez *Nicotiana benthamiana* démontrent que les 2 proteines TaICE activent la transcription de *TaCBFIVd-B9*. La structure complexe des éléments MYC dans le promoteur *CBF* du blé et l'affinité de TaICE87 , ainsi que la différence d'affinité entre TaICE87 et TaICE41 pour les éléments MYC suggèrent que ces 2 protéines peuvent différentiellement activer CBF chez le blé. De plus, la surexpression de *TaICE87* et *TaICE41* chez *Arabidopsis* entraine l'augmentation de l'expression de *AtCBF2*, *AtCBF3* ainsi que celle de plusieurs gènes régulés par le froid, et par conséquent augmente la tolérance au gel. Ces résultats suggèrent que *TaICE87* and *TaICE41* sont des homolgues fonctionnels *d'AtICE.*

Mots clés: *Arabidopsis*, liaison a l'ADN, tolérance au gel, promoteur, facteur de transcription, transactivation.

Abstract

Two different inducers of CBF expression (ICE1-like genes), *TaICE87* and *TaICE41*, were isolated from wheat. Both genes are expressed constitutively in spring and winter wheats and their encoded proteins share 94% homology within the bHLH domain. Gel mobility shift analysis showed that TaICE87 and TaICE41 can bind differentially to MYC elements in the wheat *TaCBFIVd-B9* promoter. Transient expression assays in *Nicotiana benthamiana*, showed that both TaICE proteins can activate *TaCBFIVd-B9* transcription. The complex structure of MYC elements in wheat CBF promoters and the different affinities of TaICE87 and TaICE41 for MYC elements suggest that these proteins might activate CBF differentially in wheat. Over-expression of *TaICE87* and *TaICE41* in *Arabidopsis* increased the expression of *AtCBF2*, *AtCBF3*, several cold-regulated genes and enhanced tolerance to freezing stress. These results suggest that *TaICE87* and *TaICE41* are functional homologs of *AtICE*.

Keywords: *Arabidopsis*, DNA binding, freezing tolerance, promoter, transcription factor, transactivation.

Introduction

Plants have developed strong mechanisms to adapt their cellular metabolism to survive under different stress conditions by undergoing several physiological, biochemical and molecular changes. During cold acclimation, temperate plants increase freezing tolerance (FT) by prior exposure to low, non-freezing temperatures (Guy, 1990; Thomashow, 1999). During this period of low temperature exposure, the presence of several cold-regulated genes is correlated with the development of freezing tolerance (Knight et al., 1999; Thomashow, 1999). The promoters of many cold and dehydration-responsive genes in A*rabidopsis* have been shown to contain the C-Repeat (CRT)/Dehydration Responsive Element (DRE) (Baker et al., 1994; Yamaguchi-Shinozaki and Shinozaki, 1994). The DRE/CRT related motifs have also been reported as low-temperature-responsive elements (LTREs) in the promoters of the *Brassica napus BN115* (Jiang et al., 1996) and the wheat *WCS120* (Ouellet et al., 1998) genes. C-repeat/dehydration-responsive element Binding Factor (CBF)/Drought Responsive Element Binding (DREB) proteins are the most characterized transcription factors that bind to the CRT/DRE element and activate cold- and/or drought-responsive gene expression (Stockinger et al., 1997; Gilmour et al., 1998; Liu et al., 1998; Shinwari et al., 1998; Thomashow, 2001). These transcripts encode transcriptional activators that are members of the Apetala2 (AP2)/Ethylene Responsive Element Binding Protein (EREBP) family of DNA-binding proteins (Riechmann and Meyerowitz, 1998). In *Arabidopsis*, three CBF transcripts designated *CBF1*, *CBF2* and *CBF3* (Jaglo-Ottosen et al., 1998; Medina et al., 1999), or *DREB1b*, *DREB1c* and *DREB1a* respectively (Liu et al., 1998) start to accumulate within 15 min of plant exposure to low temperature. Constitutive over-expression of the *CBF1/DREB1b*, *CBF2/DREB1c* or *CBF3/DREB1a* (Gilmour et al 2004; Jaglo-Ottosen et al., 1998; Kasuga et al., 1999; Liu et al., 1998) genes in

transgenic *Arabidopsis* plants induces the expression of multiple cold-responsive CRT/DRE-containing genes without a low temperature stimulus and confers improved freezing tolerance, indicating that these transcription factors have redundant functions and are sufficient for cold acclimation (Gilmour et al., 2000; 2004). This explains why the *CBF/DREB1* genes are themselves cold-regulated (Gilmour et al., 1998; Liu et al., 1998; Shinwari et al., 1998). However, the mechanism of *CBF/DREB1* gene induction in response to low temperature stimuli is not completely understood (Shinwari et al., 1998).

Several reports identified direct regulators of CBF/DREB1 expression, including MYB15, ICE1, and HOS1 (Agarwal et al., 2006; Dong et al., 2006; Chinnusamy et al., 2006). MYB15 binds to CBF/DREB1 promoter regions to repress its expression and negatively regulate freezing tolerance (Agarwal et al., 2006). ICE1 (for inducer of CBF/DREB1 expression) physically interacts with the Myb domain protein 15 (MYB15). ICE1 may attenuate directly (through binding to MYB15 promoter) or indirectly (through its downstream genes) MYB15 expression in response to cold (Agarwal et al., 2006). ICE1 is a MYC-like basic helix-loop-helix transcription factor that activates CBF/DREB1 expression in response to low temperature (Chinnusamy et al., 2003). ICE1 binds to canonical MYC *cis*-elements (CANNTG) in the CBF3/DREB1A promoter to induce its expression, which then leads to the induction of the CBF/DREB1 regulon (Chinnusamy et al., 2003; Lee et al., 2005). ICE1 is an important controller of CBF3/DREB1A, of the CBF/DREB1 regulon gene expression, and of freezing tolerance responses (Chinnusamy et al., 2003). Gilmour et al. (1998) hypothesized that ICE is present in an inactive state at normal growth temperature and becomes active upon exposure to low temperature to stimulate transcription of the *CBF/DREB1* genes. The high expression of osmotically responsive genes 1 (HOS1) is a RING-type ubiquitin E3 ligase that negatively regulates CBF/DREB1 expression (Ishitani et al., 1998; Dong et al., 2006). HOS1 migrates to the nucleus in response to cold treatment and polyubiquitinates ICE1, targeting this transcription factor for proteasome degradation (Lee et al., 2001; Dong

et al., 2006). Together, these results indicate that the ubiquitin E3 ligase HOS1 and MYB15 act as negative regulators while the MYC transcription factor ICE1 acts as a positive regulator to modulate expression of *CBF3/DREB1A*. Los 4 encodes a DEAD-box RNA helicase that appears to have a positive effect on CBF expression since los-4 mutant plants are chilling-sensitive. Furthermore, ectopic expression of CBF3 in these mutants was shown to restore the normal phenotype (Gong et al., 2002; Gong et al., 2005). On the other hand, the mutation of ICE1, ice1 resulted in complete elimination of *CBF3* transcript accumulation (Chinnusamy et al., 2003). Different MYC elements potentially involved in the cold response were also identified in the *CBF2* promoter (Zarka et al., 2003). The consensus sequence CACATG was suggested as a potential binding site for the ICE1 protein. However, the ice1 mutation has little effect on *CBF2* transcript accumulation under low temperature indicating that cold regulation of the different *CBF/DREB1* gene family members is independent. This suggests that another *ICE1*-like basic Helix-Loop-Helix (bHLH) protein is involved in the regulation of *CBF2* gene expression at low temperature.

In wheat, Badawi et al. (2007) have identified several *CBF* genes that are subdivided into 10 different groups. Several CBF groups are amplified only in Pooideae and most of these amplified groups (5 out of 6) are expressed at higher levels in freezing tolerant cultivars. These observations raise the question whether *ICE*-like genes regulate CBF in wheat as shown for *Arabidopsis*. To understand this regulation and identify the genes involved in the cold acclimation pathway in wheat, we pursued molecular analysis of bHLH genes found in wheat EST genomic resources. Expression studies, bioinformatics analysis, phylogenetic and genetic studies revealed that two *ICE* genes, named *TaICE87* and *TaICE41*, encode potential activators of the low-temperature stress tolerance pathway in hexaploid wheat. DNA/protein interactions studies *in vitro* and *in vivo* demonstrated that ICE genes bind to *CBF* promoters and activate their transcription. The putative function of these genes in the control of freezing tolerance in wheat is discussed.

Materials and methods

Plant material and growth conditions

Two varieties of hexaploid wheat (*Triticum aestivum* L.) including a spring habit cultivar (Manitou) and a winter habit cultivar (Norstar) were grown in environmentally-controlled growth chambers as previously described (Danyluk et al, 2003).

Cloning of cDNAs and promoters isolation

cDNA was isolated from winter wheat Norstar as described previously (Danyluk et al, 2003; Kane et al, 2005). For promoter isolation, several genomic *T. aestivum* libraries were constructed using the Genome Walker Kit (CLONTECH, Palo Alto, Calif.) according to the manufacturer's instruction. The *TaCBFIVd-B9* promoter region was PCR-amplified from genomic DNA of wheat cultivar using the following primers: 5'-ACGTCCACATGTAGCTCTGGAGGGACAACG-3' with the AP1 primer followed by a second nested PCR using 5'-GGAGATGCTCGTGAGAGGTGTGGAGTGGT-3' with the AP2 primer. The gene-specific primers were used together with adapter primers (AP) to amplify *TaCBFIVd-B9* promoter fragments from the *T. aestivum* libraries. AP1, AP2, and adapter primers were supplied in the Genome Walker Kit. Promoter sequences were then analyzed using Plant-CARE (Lescot et al, 2002), Genomatix Promoter Database (http://www.genomatix.de/products/GPD/index.html), and PLACE (Higo et al, 1999)

Search for ICE1 sequences and phylogenetic analysis

The protein sequence of TaICE87, TaICE41 and *Arabidopsis* ICE1 were used as query sequences for TBLASTN analyses to collect *ICE1*-like genes in plants. The survey was conducted against the GenBank (http://www.ncbi.nlm.nih.gov/BLAST/), the TIGR (http://compbio.dfci.harvard.edu/tgi/plant.html), the Rice Annotation Project DataBase (http://rapdb.lab.nig.ac.jp/index.html), the JGI *Physcomitrella*

patens subsp patens v1.1 (http://genome.jgi-psf.org/Phypa1_1/Phypa1_1.home.html), and the JGI *Populus trichocarpa* v1.1 (http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html) databases. Genes and accession numbers are listed in Table S1. The sequences were translated and protein sequences were aligned with ClustalW. Character-based parsimony analysis was used for phylogenetic analyses in PAUP 4.0 (Swofford 2003). Since *ICE1* in *Arabidopsis* belongs to the bHLH domain family, three non-ICE bHLH genes from *Arabidopsis*, poplar and rice were used as outgroup (Table S1).

Electrophoresis mobility shift assays (EMSA)

The full length *TaICE87* and *TaICE41* coding regions were cloned in the pENTR4 vector then transferred to the pDEST15 vector by recombination using the Gateway technology (Invitrogen; for primers, see Table S4). The resulting plasmids pDEST15-TaICE87 and pDEST15-TaICE41 were independently transformed into *Escherichia coli* BL21-A1 to express GST-fusion proteins. The proteins were purified on GST·Bind affinity resin (Novagen) and used in EMSA to determine their binding affinity towards the *TaCBF* promoter MYC elements. Sense and complementary oligonucleotides (Fig. 4A) corresponding to MYC elements were annealed and radiolabelled with [γ-^{32}P]-ATP (Amersham) to generate double-stranded probes. DNA binding reactions were performed in a total volume of 20 µl of buffer (10 mM Tris-HCl pH 7.5, 4% glycerol, 20 mM KCl, 20 mM dithiothreitol) containing 1 µM of polydI.dC, 0.2% (v/v) Triton X-100, 2 ng (~5X10^4 CPM) of probe, and 100 ng of the recombinant GST-tagged proteins. The binding specificity was assessed by competition with a 50, 100 and 500-fold excess of unlabelled double-stranded oligonucleotides. Binding reaction mixtures were incubated for 15 min at room temperature and then resolved by electrophoresis on a 4% nondenaturing polyacrylamide gel, prepared in 0.5X TBE, at 100 V for 90 min. The gels were dried, exposed to K screens and the signal was detected with a Personal Molecular Imager FX System (Bio-Rad).

Transient expression by Agroinfiltration of *Nicotiana benthamiana*

Three binary constructs based on pBin19 or pGreenII0029 were generated for transient expression assays. The primers used for cloning are indicated in Table S4. The CBFp:GFP reporter plasmid contained the mGFPer gene under the control of the *TaCBFIVd-B9* promoter while the effector constructs contained the *TaICE41*or *TaICE87* genes under the control of the CaMV35S promoter (35S:*TaICE87* and 35S:*TaICE41*). The plasmids were independently transformed into *Agrobacterium tumefaciens* strain EHA105. The transformed Agrobacteria were used individually or in combination to infiltrate intact leaves of *Nicotiana benthamiana* as described by Kane et al., (2007). After infiltration, plants were kept at 24°C for a 3 day recovery period. Plants were then maintained at 24°C or exposed at 4°C for an additional two days (cold treatment). GFP expression was evaluated using a UV-hand lamp (Long wave ultraviolet lamp, Model BlooAp MDSK, Blak-RAY) or using a confocal system (Bio-Rad MRC1024) with a Nikon Eclipse TE300 inverted microscope, and analyzed using the LaserSharp software (Bio-Rad). The accumulation of GFP was confirmed using an anti-GFP antibody (Clontech).

Over-expression of *TaICE87* and *TaICE41* in *Arabidopsis*

The 35S:*TaICE87 and* 35S:*TaICE41* constructs (Table S2) were generated for the transformation of *Arabidopsis thaliana* ecotype Columbia by floral-dipping (Clough and Bent, 1998). Transformants were selected on medium containing MS salts and vitamins supplemented with 50 mg/L kanamycin, and resistant T2 seedlings were transferred to soil and grown to seed under LD conditions (16h photoperiod) at 24/20°C (day/night). Wild-type Columbia and plants transformed with pBIN19 (35S:GUS) were used as controls.

Molecular analyses

For nucleic acids analyses, genomic DNA and total RNA were isolated from aerial parts (wheat), or leaves (*Arabidopsis*) using DNAzol or TRIzol reagents following the manufacturer's instructions (Invitrogen). For Northern blot analysis, 10μg of total RNA were separated on formaldehyde–agarose gels. Transfers to positively charged nylon membranes and hybridizations with ^{32}P-labeled probes were performed using standard molecular biology techniques. Probes for the different *CBF*s and *COR* genes were amplified using specific primers (Table S4) to avoid cross-hybridization.

Determination of freezing tolerance

A Caltec Scientific Ltd. Model 8-792 Large Capacity Temperature Stress Chamber was used to perform the freezing tolerance tests. This instrument consists of four major component systems: a Sanyo Model MDF-792 24.75 ft^3 capacity ultra-low temperature chest freezer, a custom designed stainless steel plenum box with its integral blower and heater (provides air circulation and heating) and an Omega Engineering Inc. Model CN3002 programmable profile controller (monitors the test-chamber air temperature). The controlled action of the heater combines with the constant cooling of the freezer to achieve the desired temperature at any given time.

Non-acclimated (NA) and cold-acclimated (CA) soil-grown plants (3 weeks-old) were subjected to the following freezing treatment. The temperature was lowered gradually (2°C h^{-1}) to -6.5°C for NA or -10.5°C for CA plants and maintained at this temperature for 6 hr. The temperature was then gradually increased (2°C h^{-1}) to 4°C. To determine temperature variability in the freezer, temperatures were monitored by 4 independent thermocoupled T probes distributed in the freezer and connected to an Agilent 3497-0A data acquisition/switch unit. Freezing regimes that showed more than 0.5 °C discrepancies between the different probes were rejected. To minimize light stress effects after the freezing treatment, plants were thawed at 4°C for 24 hr in the dark and away from direct light in the growth chamber (20°C) for an additional 24 hr before returning to normal light conditions. Representative pictures were taken 2

weeks after the freezing treatment. Eighteen plants were frozen per line per assay, and the experiment was repeated 3 times.

Results

Identification and characterization of *wheat ICE* genes

Data mining of public databases resulted in the identification of several *Arabidopsis* ICE1 homologues from different species including two genes from rice (AK109915 and NM_001074519). Since rice is a close relative of wheat, the rice *ICE1*-like genes were used as query to identify wheat ESTs. Four wheat expressed sequenced tags (ESTs) (CD900164, BE422944, CA714228, and BJ260527) were used to design forward and reverse primers to clone wheat *ICE1*-like genes from a cDNA library prepared from wheat aerial tissues exposed to short periods of cold-acclimation (Houde et al., 2006). We amplified two different ICE1-like genes designated *TaICE41* and *TaICE87*. The *TaICE41* cDNA encodes a protein of 381 amino acids with a predicted molecular mass of 39.5 kDa while *TaICE87* encodes a protein of 443 amino acids with a predicted molecular mass of 46.5 kDa. The complete TaICE41 and TaICE87 proteins share 46% identity and have 50% and 47% identity with AtICE1, respectively.

Alignment of TaICE41 and TaICE87 with ICE1 proteins from *Arabidopsis* and the closest homologues from rice and poplar shows that ICE1-like proteins share highly conserved regions in the bHLH domain and in their C-terminal region (Fig. 1). A short region of 11 amino acids (Box I in Fig. 1) is conserved between TaICE87, AtICE1, CbICE53 and Os11g0523700 suggesting that these amino acids may be part of an important domain for ICE activity. The *ice1* mutant causing a loss of ICE function was found to have a substitution in this region where R236 is substituted by H236 in *Arabidopsis* (Chinnusamy et al., 2003). The box I is absent from the wheat TaICE47 protein and from a rice homolog (Os01g0928000 in Fig. 1) suggesting that this may represent an ICE gene with distinct properties. Although the N-termini of ICE1-like proteins were distinct among plant species, a weakly conserved domain was found preceding the bHLH domain (between box I and box II, Fig. 1).

A total of 28 ICE1 homolog sequences were collected from 17 plant species (Tables S1 and S2). The highly conserved region from the bHLH domain to the C-terminal region was subjected to phylogenetic analysis using the closest non-ICE bHLH proteins from *Arabidopsis*, rice and poplar as outgroup genes. The different *ICE* genes are mainly separated into three major clades (Fig. 2). ICE-like proteins in dicots form one clade while the *ICE*-like proteins from 5 monocots species, (barley, maize, rice, sugarcane and wheat) form two distinct clades. Although 133 genes were identified as bHLH genes in *Arabidopsis* (Heim et al. 2003), only AtICE1 and AtbHLH033 are included in the dicot ICE1 clade (Fig. 2). AtbHLH061 is the closest homologue to ICE1 among the remaining *Arabidopsis* bHLH genes, but it is clearly separated from *ICE1* genes.

So far, at least two ICE1-like genes are localized on distinct chromosomes in *Arabidopsis*, barley, poplar, and rice. The ICE1-like pair of proteins in dicots, (i.e. AtICE1/AtbHLH033 in *Arabidopsis* and PtrLGXII1027/PtrLGXV000793 in poplar) show high homology especially in the C-terminal region (89% and 97%, respectively). TaICE87 forms a separate clade with several monocot genes and is named ICE1-like because it is closer to the dicot ICE1 clade. TaICE41 forms a distinct clade with other monocot genes and is named ICE2-like. HvICE2 is among this ICE2-like clade and was identified as a monocot homologue to AtICE1 that is closely related but distinct from HvICE1 (Tondelli et al., 2006; Skinner et al., 2006). Two other wheat ESTs besides TaICE87 are included in the ICE1-like clade and represent distinct ICE1-like genes since their protein sequences possess 74 to 90% identity between each other and contain one or two gaps in their alignment (data not shown).

Real-time PCR quantification of mRNA expression levels of *TaICE87* and *TaICE41* during cold acclimation was performed using gene specific-primers and TaqMan probes spanning the exon-intron junction. The pattern of mRNA expression was not significantly different during low temperature (LT) treatment in both winter and spring wheat, demonstrating that *ICE1* mRNA expression is constitutive (data not shown). Since the expression of these two wheat *ICE* genes is not genotype

dependent, it indicates that the difference in freezing tolerance between cultivars is not correlated with *ICE* gene transcript levels.

CBF promoter analysis and binding specificity

Analysis of the *TaCBF* genes showed the presence of MYC recognition sequences in their promoters (Fig. 3A). The motif search revealed that the MYC binding sites in wheat and *Arabidopsis CBF* promoters contain all the 16 different possible MYC permutation sequences of the CANNTG core (Table S3). We have subdivided the different MYC elements into 4 groups according to the 3rd base of the CANNTG core sequence (MYC1 through MYC4). Each group is then subdivided into subgroups according to the 4th base (namely for MYC1: MYC1a, MYC1c, MYC1g, and MYC1t). Based on this classification, differences in the composition of MYC variants between *CBF* genes in *Arabidopsis* and wheat were observed. The *AtDDF1*, *AtDDF2* and *AtCBF4* are member of the *DREB-A1* subfamily of transcription factors in *Arabidopsis* but are not responsive to low temperature (Haake et al., 2002). The MYC2 elements are found in two or three copies in the promoter of cold-responsive CBFs (highlighted in blue in Table S3) but are generally absent from the genes that do not respond to low temperature (except for *AtCBF4* which contain one copy of a MYC2 variant). ICE1 from *Arabidopsis* was shown to bind to the MYC2a sequence in the *AtCBF3* promoter and to activate its transcription in transgenic plants (Chinnusamy et al., 2003). Analysis of 19 wheat *CBF* promoter regions revealed that the MYC2a element is present in several promoters, suggesting that it is conserved and plays a role in both wheat and *Arabidopsis*. MYC3 and MYC4g elements are generally absent in *Arabidopsis* (except MYC3a in *AtDDF1*) but found in several wheat promoters. Most interestingly, the MYC4g element is present in most of group IV *CBF* promoters suggesting that this element may have evolved specifically in wheat. MYC3 elements are found in all wheat CBF groups and thus are not likely to confer specificity between wheat CBF subgroups. Furthermore, since the MYC2

element is underrepresented in the group IV *CBF* genes, this group of genes may be regulated by factors that bind more specifically to MYC4 elements.

Among the isolated wheat promoters, the *TaCBFIVd-B9* promoter is the only one that contains both a MYC2a element and a MYC4g element. Putative cis-acting elements were searched from the 1544 bp of *TaCBFIVd-B9* promoter. A number of conserved motifs found in several eukaryotic promoters for gene expression and regulation was identified (Fig. 3B). A typical TATA box was recognized at position −158 bp (TTTATATA) upstream of the ATG translation initiation codon. Potential regulatory elements associated with hormone- and stress-related responses were located within the TaCBFIVd-B9 promoter region: 5 MYB binding sites, 14 MYC binding sites, 2 low temperature responsive element (LTRE), one ABRE-like, one CArG-box binding site for MADS transcription factors and one MeJA-like element.

The AtICE1 protein was shown to bind to a consensus bHLH core sequence (CANNTG) or E-Box sequence within the *AtCBF3* promoter (Toledo-Ortiz et al., 2003). The preferred MYC element was designated as MYC2 (CACATG, or MYC2a in Table S2). Since the MYC2a element is present in the promoter of several wheat CBF genes, this element was used for electrophoretic mobility shift assay (EMSA). MYC4g is also targeted for the assay because it is specifically found in the wheat CBF Group IV. Among 14 MYC-recognition core elements that were found within the 1544 bp of *TaCBFIVd-B9* promoter (Fig. 3B and Table S3), one MYC2a, two MYC4g and 11 other MYC elements were identified. We hypothesized that the MYC2a and MYC4g elements may be bound with different specificities by the two isolated TaICE proteins. The MYC2a and the two different MYC4g sequences (MYC4g1 and MYC4g2, in Fig. 3B) were used to design DNA probes for EMSA studies (Fig. 4). Purified TaICE87 did not bind to MYC2a (Fig. 4B panel I) while it bound to MYC4g1 and MYC4g2 elements (Fig. 4B panels II and III). Competition with unlabeled MYC4g (1 or 2) completely inhibited the binding of TaICE87 to its targets (Fig. 4B panels II and III). On the contrary, TaICE41 bound specifically to MYC2a (Fig. 4C panel I) but not to MYC4g1 or MYC4g2 (Fig. 4C panel II and III).

Taken together, TaICE87 and TaICE41 have preferences in binding to MYC elements. This differential affinity gives these trans-acting elements the potential to regulate the transcription of different *TaCBF* genes.

TaICE87 and TaICE41 transactivate the wheat *TaCBFIVd-B9* promoter

The ability of the two wheat ICE proteins, to activate the wheat *TaCBFIVd-B9* promoter was determined using the transient expression system based on the Agroinfiltration of *Nicotiana benthamiana* leaves. The plasmid pGreenII0029 containing the CaMV35S promoter (PGII) was used to over-express TaICE87 or TaICE41, while a reporter plasmid was constructed to express the Green Fluorescent Protein (GFP) under the control of the *TaCBFIVd-B9* promoter (CBFp) (Fig. 5A). Control experiments with non-infected plants or plants infected with the PGII vector showed no green fluorescence using a UV-hand lamp (Fig. 5B, panel 1) or using confocal microscopy (Fig. 5B, panels 2 and 3). Similarly, plants infected with CBFp and treated at 24°C did not show any GFP fluorescence in whole plants (Fig. 5B, panel 4) or under the microscope (Fig. 5B, panels 5 and 6). However, plants infected with CBFp and treated at 4°C showed no visible fluorescence under the UV-hand lamp (not shown) but a weak GFP-fluorescence was observed under the microscope indicating that tobacco can respond to low temperature and activate the wheat CBF promoter (Fig. 5B, panel 6). Interestingly, when plants were co-infected with CBFp and either of the effector genes (*TaICE87* and *TaICE41*), a strong GFP expression was clearly seen in the whole leaf using a UV-hand lamp (Fig. 5B, panel 7) and a strong fluorescence is seen in epidermal cells under confocal microscopy at 24°C (Fig. 5B, panels 8 and 9) or at 4°C (data not shown). These results demonstrate that both wheat ICE genes strongly transactivate the *TaCBFIVd-B9* promoter in *Nicotiana benthamiana*. Immunoblot analysis of GFP expression showed that TaICE87 was more efficient in transactivating the CBP promoter compared to TaICE41 (Fig. 5C). Taken together, these results showed that both TaICE87 and TaICE41 proteins function as transcription activators of the *TaCBFIVd-B9* promoter.

TaICE41* and *TaICE87* over-expression enhances FT in *Arabidopsis

The function of wheat *ICE* genes were assessed in *Arabidopsis* by analyzing the expression level of *AtCBF1, AtCBF2*, *AtCBF3*, and different cold-regulated (*COR*) genes in three independent transgenic lines showing an increased expression of *TaICE87* or *TaICE41*. Their effect on FT were evaluated in the non-acclimated (NA) and cold-acclimated (CA) transgenic and control plants exposed to a freezing temperature of -6.5°C or -10.5°C, respectively. Plant survival was scored after 2 weeks of recovery under normal growth conditions, and representative results are shown in Fig. 6A and 6B for *TaICE87* and *TaICE41*, respectively. NA plants showed no difference in freezing tolerance between control plants, including the wild type (WT) and empty vector transformed plants, and both transgenic lines over-expressing *TaICE87* or *TaICE41* genes. However, the over-expression of *TaICE87* or *TaICE41* genes resulted in increased FT after cold acclimation compared to the WT or empty vector (Fig. 6C and 6D).

The over-expression of *TaICE87* or *TaICE41* did not modify the expression of *AtCBF1* (data not shown). The over-expression of *TaICE87* resulted in the greater accumulation of the cold responsive genes, *AtCBF2*, *AtCBF3*, *COR6.6*, and *COR15a*, compared to WT or the empty vector (Fig. 7A and 7B). The over-expression of *TaICE41* increased the expression of *AtCBF2*, *AtCBF3* and *AtCOR6.6* genes but not of *AtCOR15a* compared to WT or vector alone. The difference in the level of gene expression and specificity between *TaICE41* and *TaICE87* suggests that *TaICE87* and *TaICE41* are functional homologs of *AtICE* genes with different binding affinities. Alternatively, the difference in the level of *TaICE* RNA (Fig. 7) expression may be in part responsible for the difference in the level of over-expression of the target genes.

Discussion

We isolated and characterized two functional wheat ICE-like bHLH genes that are candidate regulators of CBF gene expression and freezing tolerance in wheat. Sequence comparison showed that TaICE87 and TaICE41 shared high amino acid identity with *Arabidopsis*, ICE1 transcription factors especially in the bHLH domain and the C-terminal region. Phylogenetic analysis of 28 ICE1 homologs from monocot and dicot species revealed the presence of three major clades. Only one ICE1-like clade was found in eudicots while two ICE-like clades were identified in monocots suggesting that ICE1-like genes have diverged from the ancestral ICE-like gene into monocot and eudicots while ICE2-like genes evolved distinctively in monocots. The presence of two distinct clades in monocots suggests that these ICE proteins may have different properties. The TaICE87 protein is grouped in the ICE-like clade because it is phylogenetically closer to the eudicot ICE1 clade while the TaICE41 protein is in a distinct clade which was named ICE2-like. One of the major differences between the two ICE clades is the presence of additional amino acids toward the end of Box II (near aa 371 of AtICE1) that modifies the conserved LPPT sequence, and the absence of Box I in the ICE2-like genes (shown only for TaICE41 and OS01g092800 in Fig. 1). *TaICE87* and *TaICE41* are constitutively expressed but at a low level in wheat (data not shown). The over-expression of either *TaICE87* or *TaICE41* genes in *Arabidopsis* resulted in enhanced FT after cold acclimation only suggesting that other factors induced by LT exposure are needed to activate the wheat ICE genes and improve FT. This is consistent with the result obtained by Chinnusamy et al. (2003). Cold-induced modification of the ICE1 protein or of a transcriptional cofactor appears to be necessary for ICE1 to activate the expression of CBFs. AtICE1 binds to MYC recognition sites in the *AtCBF3* promoter to induce its expression but has a minimal effect on the expression of *AtCBF1* and *2* (Chinnusamy et al. 2003).

We have found that TaICE87 strongly activates both AtCBF2 and AtCBF3 suggesting that this ICE protein has a broader spectrum of affinities than AtICE1 to bind the MYC elements present in the *Arabidopsis* promoters. The presence of several MYC variants in these cold responsive *CBF* promoters may provide appropriate binding sites for this wheat ICE protein. Since AtICE1 was shown to regulate only *AtCBF3*, other AtICE-like protein may be involved in the regulation of one of the other cold-regulated CBF in *Arabidopsis*. The AtbHLH033 gene which is found within the eudicot ICE1 clade contains a Box I that is well conserved (except for one amino acid) but has a short deletion of 5 amino acids that covers the LPPT sequence which is well conserved between AtICE1 and TaICE87 towards the end of Box II and present in most proteins of the eudicot clade, with the exception of BnDY000939 and AtbHLH033. There is no deletion of this region in TaICE41 but there is a short addition and a less conserved sequence in this region which may be related to the weaker effect on the over-expression of *AtCBF2* and *AtCBF3* genes. Further studies will be required to characterize the role of the different protein regions in relation to the binding specificity of ICE proteins. The presence of a MYC2a element in the promoter of *AtCBF2* without activation by AtICE1 suggests that upstream and downstream sequences surrounding MYC2a might have a significant effect on the binding activity of AtICE1. Furthermore, further analyses of the MYC element composition of the *DREB-1A* promoters show that the MYC2c variant is present only in the promoter of *CBF3* (Table S3) suggesting that this variant may be important to increase the binding strength of AtICE1 or other ICE-like proteins for cold activation. The MYC2c variant has not been tested in previous studies and quantitative analysis of the binding affinity will be needed to compare the different MYC elements (Xue 2005). Similarly, there are 3 different MYC4 elements in the *AtCBF3* promoter while there is only one in *AtCBF1* and *AtCBF2*. The presence of these elements (and surrounding DNA sequence) may provide more specific binding to AtICE1. In the same manner, the greater specificity of TaICE87 to bind wheat MYC4g rather than wheat MYC2a and its ability to activate *AtCBF3* and *AtCBF2*

may indicate that MYC4 or other MYC variants may be able to recruit the wheat TaICE87 protein to activate the transcription of *AtCBF3* and *AtCBF2*. The ability of TaICE41 to bind to wheat MYC2a and its ability to activate the *AtCBF2 and AtCBF3* promoters support the need for detailed quantitative characterization of the binding affinities of various MYC elements.

Analysis of the *CBF* promoter sequences from wheat indicated that the *TmCBF* and *TaCBF* genes also contain several MYC recognition sequences in their promoters. Transient expression assays in tobacco showed that either TaICE87 or TaICE41 activate the *TaCBFIVd-B9* promoter. *In vitro* binding assays showed that TaICE87 can bind to two different oligomers containing MYC4g motifs in the *TaCBFIVd-B9* promoter while it could not bind the MYC2a element. In contrast, TaICE41 binds to MYC2a rather than MYC4g elements. This demonstrates that TaICE87 and TaICE41 can act as direct activators of *TaCBFIVd-B9* expression in distinct ways. These results suggest that at least two different ICE can activate *CBF* genes in wheat. Without the upstream gene sequence for all the *TaCBF* genes and extensive EMSA studies to test all the potential MYC binding sites, it would be premature to speculate as to the preference of TaICE87 or TaICE41 proteins for the different *TaCBF* genes. However, our survey of potential MYC recognition sites in wheat *CBF* promoters revealed that these promoters are enriched in specific MYC element subtypes supporting the potential for differential regulation of *CBF* genes by the two *TaICE* genes. In particular, the MYC2 elements are seldom found in the *TaCBF* group IV promoters while the MYC4g is found exclusively within this group. Furthermore, the MYC3 element is underrepresented in cold regulated *Arabidopsis CBF*s while they are found in most wheat *CBF* promoters suggesting that these elements may be important for wheat *CBF* transcription. Although our previous work has shown that specific *CBF* genes are much more expressed in winter compared to spring wheat (Badawi et al., 2007), it is difficult to make any correlations between the level of expression of the different *CBF* genes in winter/spring cultivars regarding the role of particular MYC sequences. Phylogenetic analysis showed that wheat Group II and

Group III *CBF* genes are closer to *AtCBF*, compared to the Group IV which amplified more recently in *Pooideae* (Badawi et al., 2007). This amplification of *CBF* genes in wheat appears to be accompanied by a specialization of *TaICE* function. The impact of this specialization increases the complexity of the regulatory network but may also be a redundant mechanism that increases the robustness of the cold acclimation process in the hardy winter wheat. The production of transgenic wheat over-expressing the different *TaICE* genes will be required to compare their ability to activate different *CBF* genes. Alternatively, the isolation of additional wheat *CBF* promoters from the different groups and the quantitative analysis of promoter efficiency (Xue 2005) using the identified MYC elements will help to evaluate the ability of the different ICE proteins to specifically bind to particular *CBF* promoters. Transactivation experiments will also be useful to document the specificity of *CBF* gene regulation. The cloning of *CBF* promoter regions from winter and spring wheat within the *CBF* groups that are more expressed in winter wheat may reveal subtle differences in MYC element composition that contribute to the differences observed in their expression (Badawi et al., 2007).

The *TaICE87* and *TaICE41* genes were also evaluated for their ability to activate *CBF* transcription and increase FT in *Arabidopsis*. Freezing tolerance of *TaICE87* and *TaICE41* over-expressing plants (Fig. 7) was greater than that of WT under cold-acclimated conditions while no significant difference was observed under non-acclimated conditions, similar to that of *AtICE1*-over-expressing plants (Chinnusamy et al., 2003; Miura et al. 2007). This over-expression of *TaICE87* or *TaICE41* in *Arabidopsis* showed that the two wheat genes are functional in *Arabidopsis* since they are able to increase *AtCBF* gene expression and improve freezing tolerance.

A potential sumoylation site was found in both TaICE41 and TaICE87 proteins using SUMOplot (Minty et al., 2000) (Fig. 1). Since the sumoylation site is conserved between AtICE1, TaICE87 and TaICE41, it suggests that TaICE41 and TaICE87 activity could be regulated by sumoylation via the SUMO E3 ligase. These findings

are consistent with the result of Miura et al. (2007), who showed that sumoylation of AtICE1 protein by SIZ1, a SUMO E3 ligase, plays a role in its activation and/or stability. AtSIZ1-mediated sumoylation of the MYC-like transcription factor ICE1, at K393, is required to activate CBF3/DREB1A-dependent cold signaling, *COR* gene induction and freezing tolerance in *Arabidopsis*.

In conclusion, we describe here the cDNA cloning and functional characterization of two ICE-like (bHLH) transcription factor genes, *TaICE87* and *TaICE41*, from wheat. Both proteins possess a conserved bHLH domain and are constitutively expressed. The DNA-binding domain of TaICE87 and TaICE41 interacts with MYC2a and MYC4g respectively in the *TaCBFIVd-B9* promoter *in vitro* and activate its transcription. Their over-expression in *Arabidopsis* activates CBF transcription and improves freezing tolerance. This is consistent with their putative role in wheat.

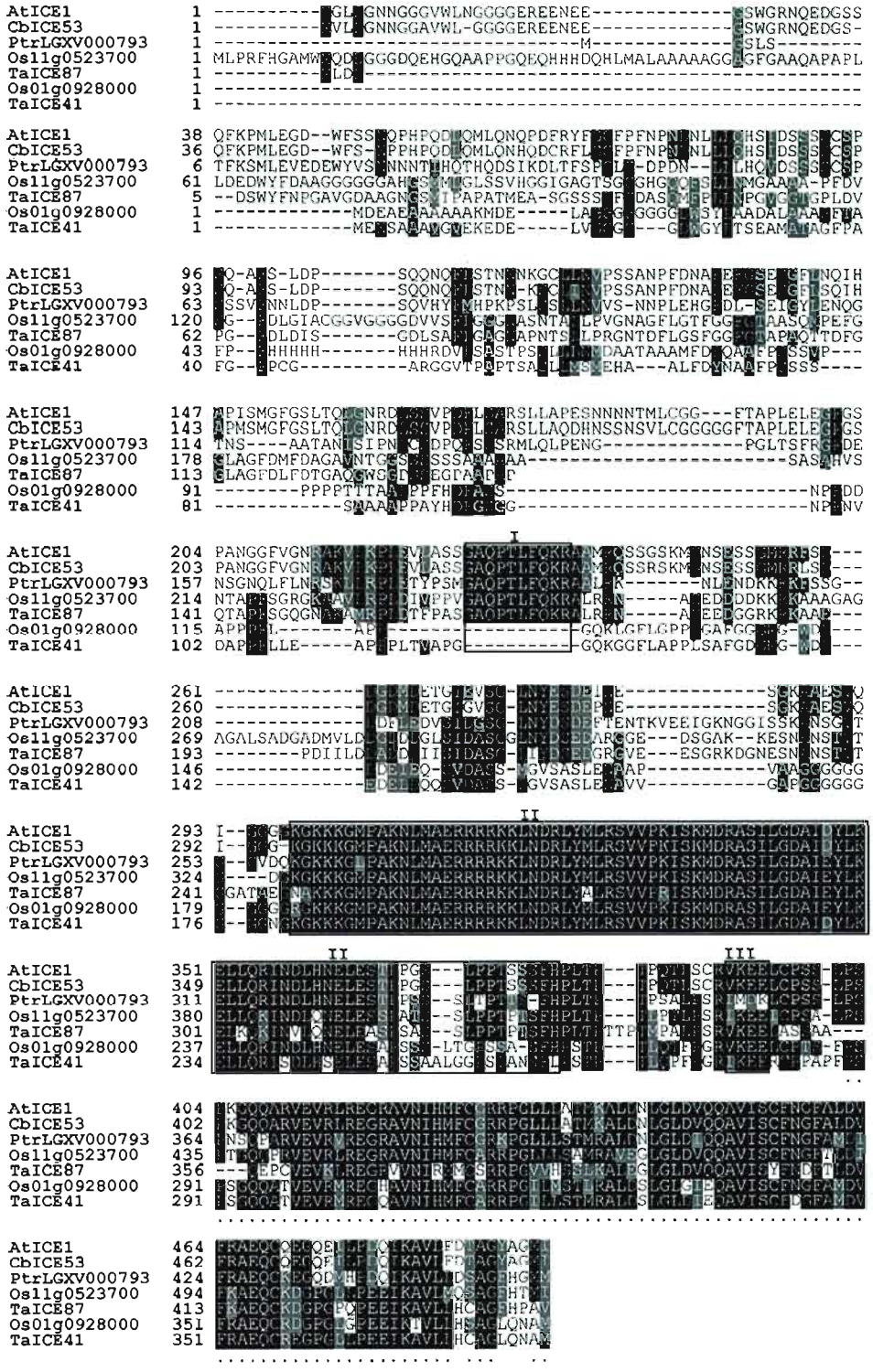
Acknowledgments

We thank K. Tremblay for her technical assistance, and all lab members for helpful discussions. This work was supported by grants from the Natural Sciences and Engineering Research Council, Genome Canada and Génome Québec (MH and FS).

Figures

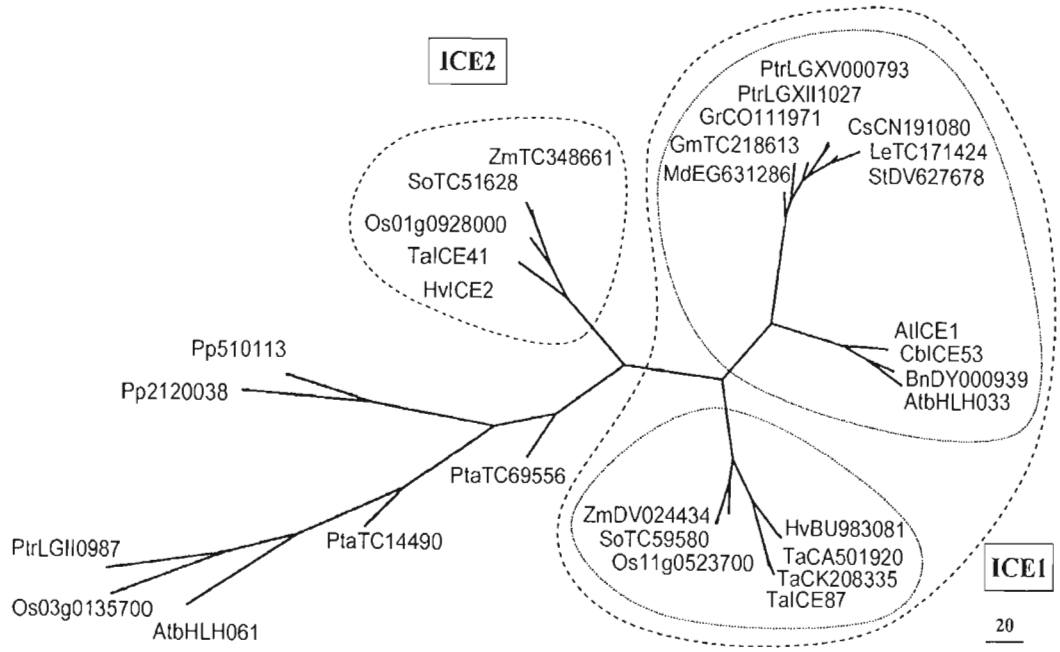
III- Figure 1: Sequence alignment and phylogenetic analysis of ICE-like proteins from different species.

The entire protein sequence of ICE1 from *Arabidopsis*, its closest homologues from *Capsella bursa-pastoris*, rice (one close homologue of *TaICE87* and one close homologue of *TaICE41*) and poplar, and wheat ICE1-like proteins were aligned using ClustalW. Identical and similar residues are highlighted in black and gray, respectively. A stretch of 11 amino acids (GAQPTLFQKRA) that is totally conserved between TaICE87, AtICE1 and CbICE53 and specifically found in *ICE1*-like genes is boxed (box I). The 56aa bHLH domains and ZIP regions are indicated by box II; and the C-terminal conserved region is indicated by dots. The SUMO conjugation motif is indicated (box III).


AtICE1
CbICE53
PtrLGXV000793
Os11g0523700
TaICE87
Os01g0928000
TaICE41
I
II
III

III- Figure 2: Phylogenetic tree of ICE1-like proteins from plants.

The maximum parsimony tree is constructed using the deduced amino acid sequences of bHLH domain and the conserved C-terminal region. The scale indicates the number of amino acid substitutions between taxons. All the *ICE1*-like genes from dicots form a single clade. TaICE87 and TaICE41 are separated to distinct clades with other *ICE* genes from monocots. The clade involving TaICE87 is clustered close to the ICE1-like clade from dicots, suggesting that genes in this clade are ICE1-like monocots genes. TaICE41 was included in another clade with monocot genes that diverged from the ICE1-like clade and was named ICE2-like. At: *Arabidopsis thaliana*, Bn: *Brassica napus*, Cb: *Capsella bursa-pastoris*, Cs: *Citrus sinensis*, Gm: *Glysine max*, Gr: *Gossypium raimondii*, Hv: *Hordeum vulgare*, Le: *Lycopersicon esculentum*, Md: *Malus domestica*, Os: *Oryza sativa*, Pp: *Physcomitrella patens*, Pta: *Pinus taeda*, Ptr: *Populus tremula*, So: *Saccharum officinarum*, St: *Solanum tuberosum*, Ta: *Triticum aestivum*, Zm: *Zea mays.*

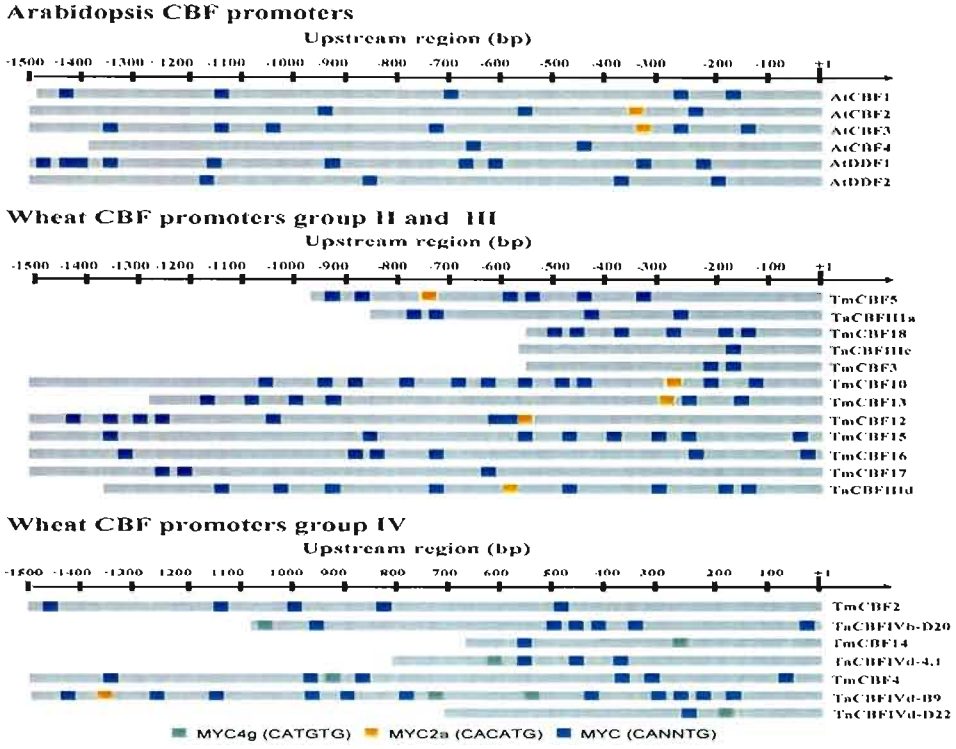

ICE2
PtrLGXV000793
PtrLGXII1027
GrCO111971
CsCN191080
GmTC218613
LeTC171424
MdEG631286
StDV627678
ZmTC348661
SoTC51628
Os01g0928000
TaICE41
HvICE2
AtICE1
CbICE53
BnDY000939
AtbHLH033
Pp510113
Pp2120038
PtaTC69556
ZmDV024434
SoTC59580
Os11g0523700
HvBU983081
TaCA501920
TaCK208335
TaICE87
ICE1
PtrLGII0987
PtaTC14490
Os03g0135700
AtbHLH061
20

III-Figure 3: Distribution of potential MYC elements in different CBF promoters and structure of the *TaCBFIVd-B9* promoter in wheat.

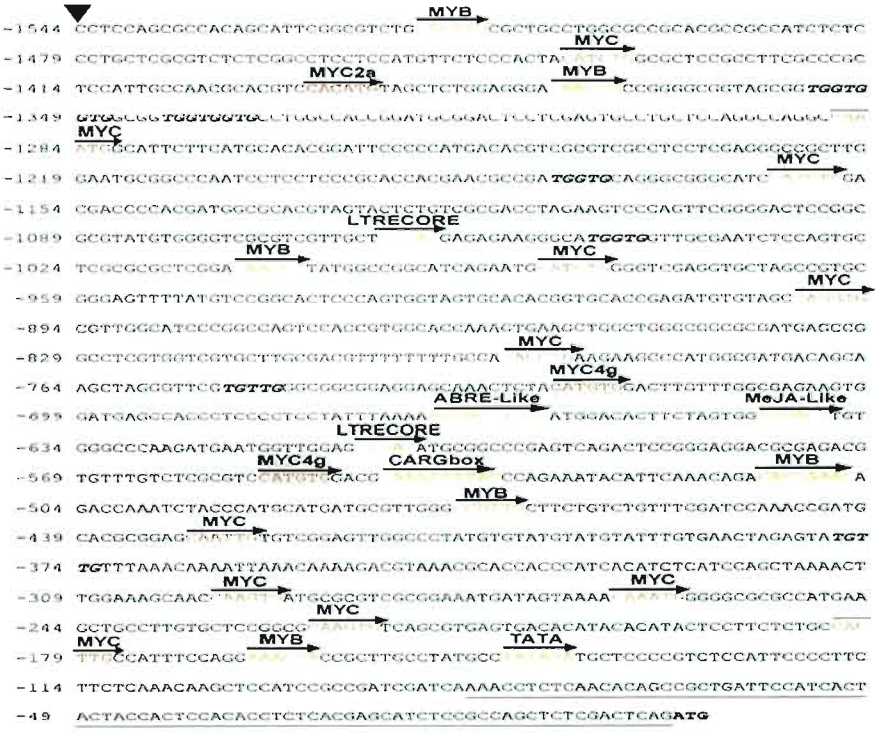
(A) Promoter maps of representative *Arabidopsis* Group-II, Group III and Group IV *CBF* promoters from wheat showing the relative position and distribution of two previously identified MYC elements (CACATG: MYC2a; and CATGTG: MYC4g).

(B) Nucleotide sequence of 5′-flanking promoter region and putative *cis*-acting elements of the *TaCBFIVd-B9* gene. The 5′ untranslated region is underlined. The plant CARE and PLACE programs were used for promoter analysis.

A


Arabidopsis CBF promoters
Upstream region (bp)
-1500 -1400 -1300 -1200 -1100 -1000 -900 -800 -700 -600 -500 -400 -300 -200 -100 +1
AtCBF1
AtCBF2
AtCBF3
AtCBF4
AtDDF1
AtDDF2
Wheat CBF promoters group II and III
Upstream region (bp)
-1500 -1400 -1300 -1200 -1100 -1000 -900 -800 -700 -600 -500 -400 -300 -200 -100 +1
TmCBF5
TaCBFIIIa
TmCBF18
TaCBFIIIc
TmCBF3
TmCBF10
TmCBF13
TmCBF12
TmCBF15
TmCBF16
TmCBF17
TaCBFIIId
Wheat CBF promoters group IV
Upstream region (bp)
-1500 -1400 -1300 -1200 -1100 -1000 -900 -800 -700 -600 -500 -400 -300 -200 -100 +1
TmCBF2
TaCBFIVb-D20
TmCBF14
TaCBFIVd-4.1
TmCBF4
TaCBFIVd-B9
TaCBFIVd-D22
MYC4g (CATGTG)
MYC2a (CACATG)
MYC (CANNTG)

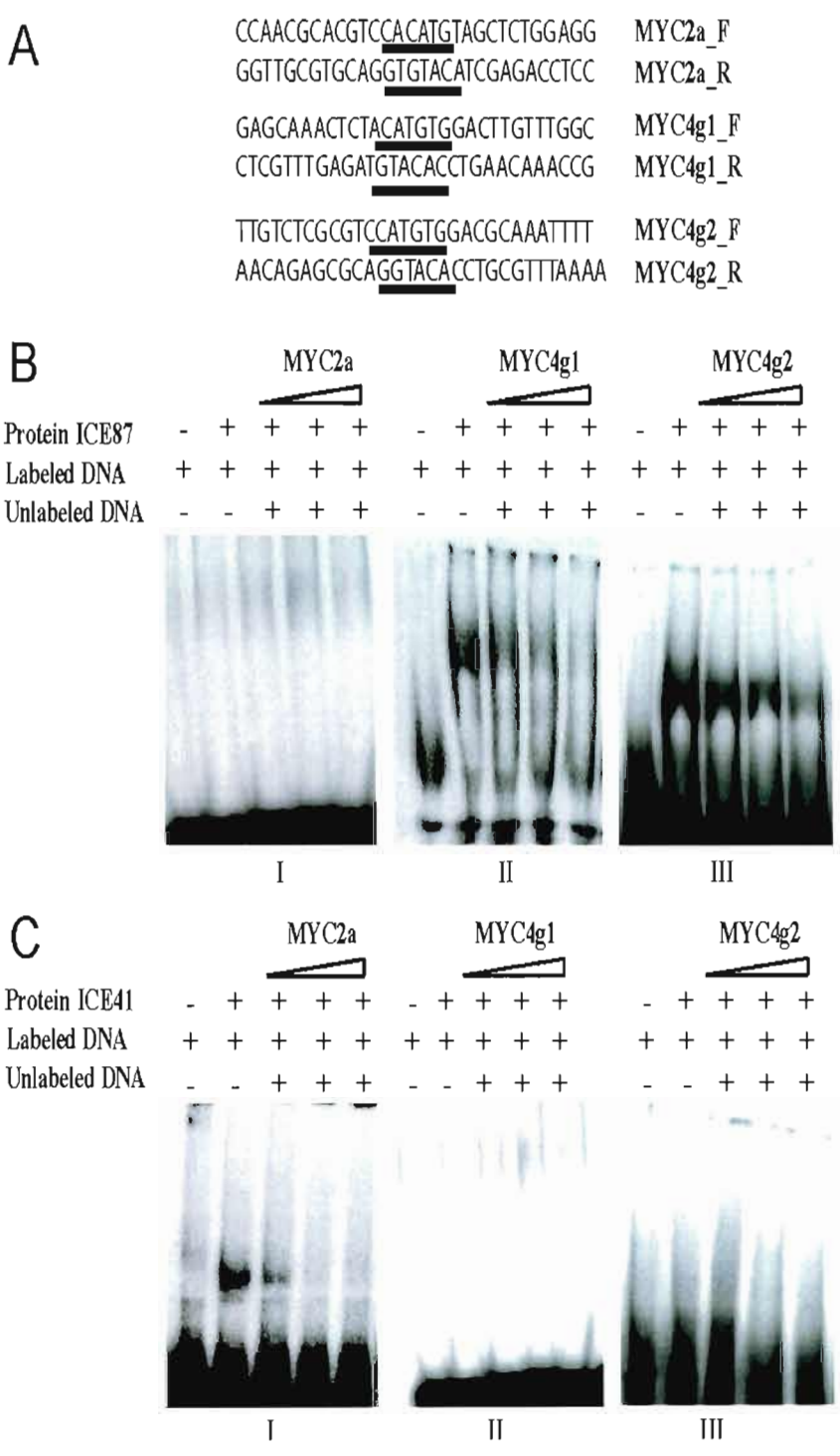
B


-1544
-1479
-1414
-1349
-1284
-1219
-1154
-1089
-1024
-959
-894
-829
-764
-699
-634
-569
-504
-439
-374
-309
-244
-179
-114
-49
MYB
MYC
MYC2a
MYB
MYC
MYC
LTRECORE
MYB
MYC
MYC
MYC
MYC4g
ABRE-Like
MeJA-Like
LTRECORE
MYC4g
CARGbox
MYB
MYB
MYC
MYC
MYC
MYC
MYB
TATA

III-Figure 4: DNA-binding affinities of the recombinant TaICE87 and TaICE41 proteins to different MYC elements of the *TaCBFIVd-B9* promoter.

(A) Double-stranded oligomers of MYC2a, MYC4g1 and MYC4g2 used as probe in the DNA-binding assays.

(B) and (C) Electrophoretic mobility shift assays. Gel mobility shift assay showing a different binding affinity between the recombinant TaICE87 (B) and TaICE41(C) proteins and the MYC2a, MYC4g1 and MYC4g2 sequences. HIS-tagged recombinant proteins were incubated with or without competitors on ice for 20 min. The P^{32} labeled probe (50,000 cpm) was then added and the mixture incubated at 25°C for 30 min. The labeled MYC fragment used in each experiment is indicated at the top of each panel. Triangles indicate increasing amounts of unlabeled competitor which corresponds to 50, 100 and 500-fold excess of each probe.

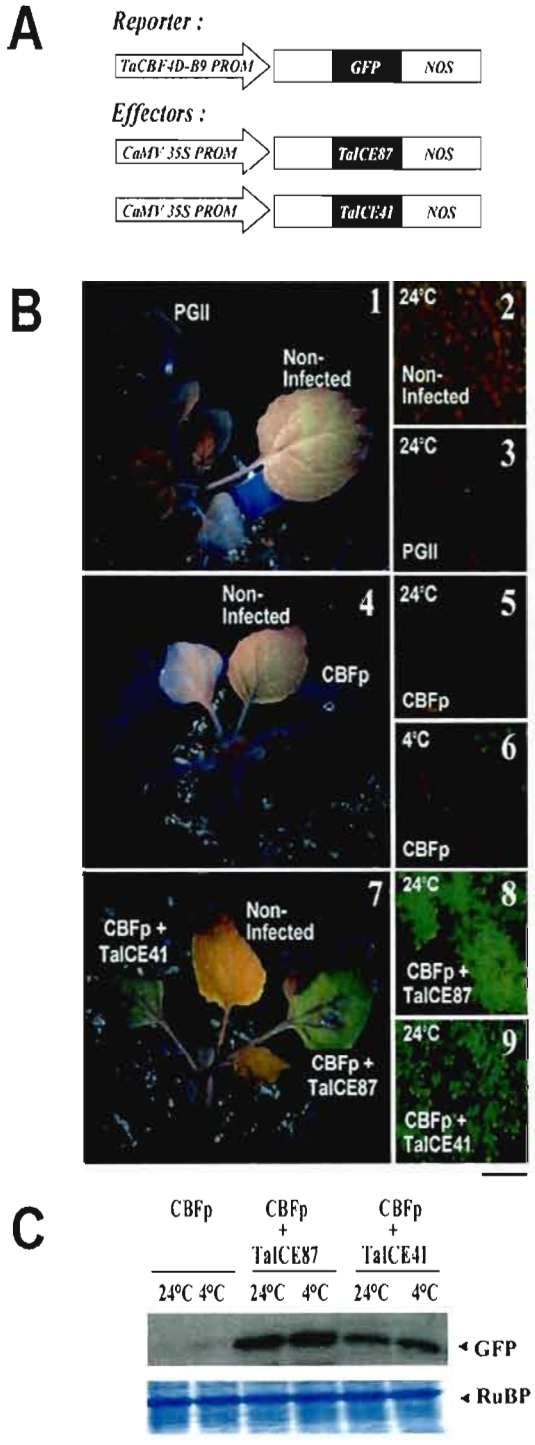

A
CCAACGCACGTCCACATGTAGCTCTGGAGG MYC2a_F
GGTTGCGTGCAGGTGTACATCGAGACCTCC MYC2a_R
GAGCAAACTCTACATGTGGACTTGTTTGGC MYC4g1_F
CTCGTTTGAGATGTACACCTGAACAAACCG MYC4g1_R
TTGTCTCGCGTCCATGTGGACGCAAATTTT MYC4g2_F
AACAGAGCGCAGGTACACCTGCGTTTAAAA MYC4g2_R
B
MYC2a
MYC4g1
MYC4g2
Protein ICE87 - + + + + - + + + + - + + + +
Labeled DNA + + + + + + + + + + + + + + +
Unlabeled DNA - - + + + - - + + + - - + + +
I
II
III
C
MYC2a
MYC4g1
MYC4g2
Protein ICE41 - + + + + - + + + + - + + + +
Labeled DNA + + + + + + + + + + + + + + +
Unlabeled DNA - - + + + - - + + + - - + + +
I
II
III

III-Figure 5: Transactivation of the *TaCBFIVd-B9* promoter-GFP fusion gene by the TaICE87 or TaICE41 proteins.

(A) Reporter and effector constructs used in the transient assays. CBFp: GFP reporter gene under the control of the wheat *TaCBF4D-B9* promoter; *TaICE87* or *TaICE41*: *TaICE87* or *TaICE41* cDNA under the control of the CaMV35S promoter; GFP: green fluorescent protein; NOS: nopaline synthase terminator.

(B) Transactivation experiments in *Nicotiana benthamiana*. Intact leaves were infiltrated with *Agrobacterium* strains carrying the CBFp construct with or without ICE effectors. Panels 1, 4 and 7 show whole plants exposed to a UV-hand lamp. Panels 2, 3, 5, 6, 8 and 9, GFP fluorescence detection by laser scanning confocal microscopy in leaf epidermal cells 7 days post-infection. Non-infected: normal leaf without *Agrobacterium* infection; PGII: leaf infiltrated with the pGreenII vector only (CaMV35S promoter alone); CBFp: reporter constructs (GFP under the control of the wheat CBF promoter); GFP + TaICE87: co-infection with the reporter and TaICE87 effector constructs; GFP + TaICE41: co-infection with the reporter and TaICE41 effector constructs. Bar: 60 µm. The data shown are representative of at least three independent experiments (n = 16 plants).

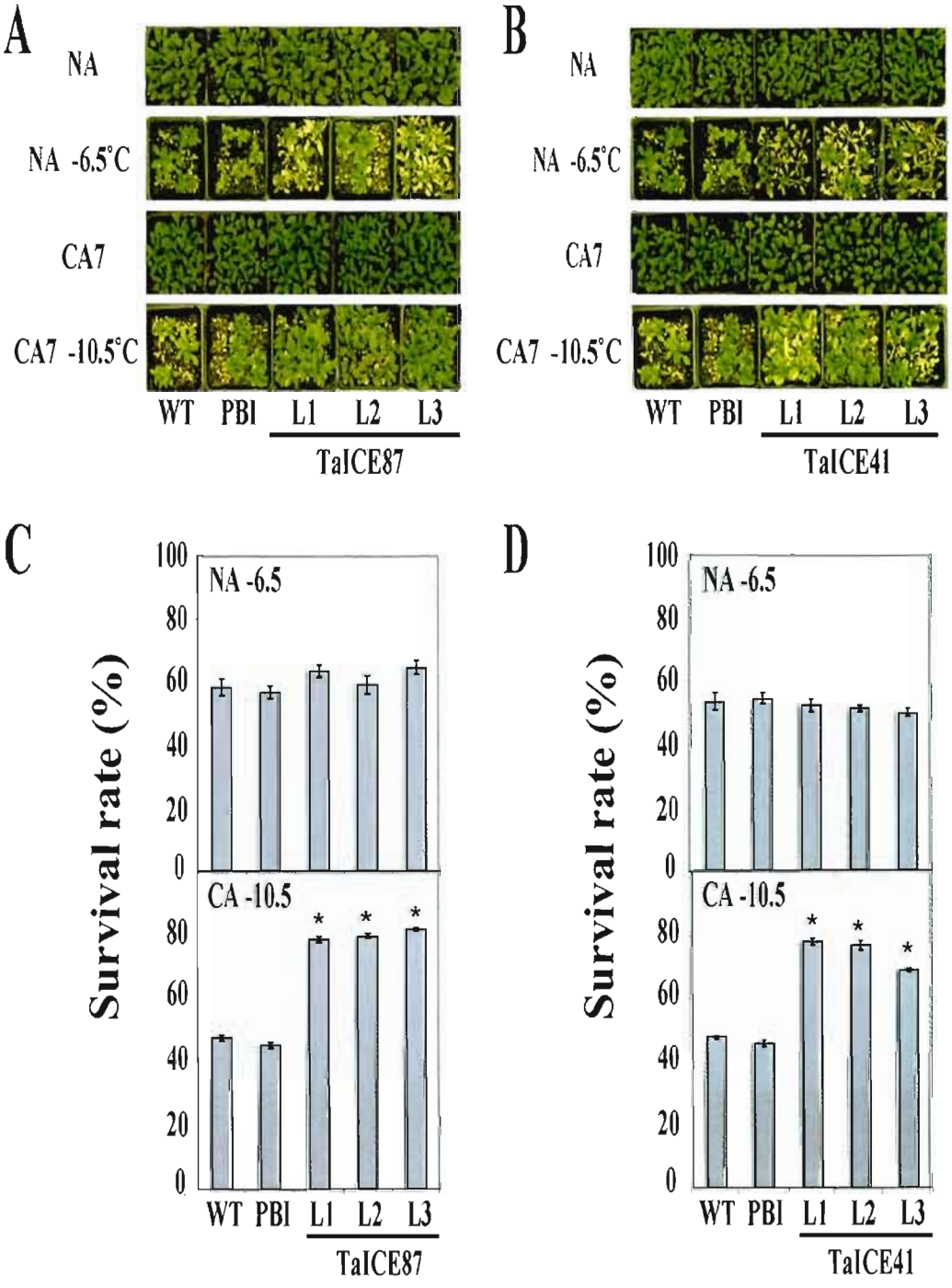
(C) Immunoblot analysis of GFP protein accumulation in the Agroinfiltrated *Nicotiana* leaves. Soluble proteins were separated by SDS-PAGE, transferred to PVDF and probed with the anti-GFP antibody. CBFp: GFP reporter construct; CBFp + TaICE87: co-infection with the reporter and TaICE87 effector constructs; CBFp + TaICE41: co-infection with the reporter and TaICE41 effector constructs. A Coomassie brilliant blue stained gel is shown as loading control. The same results were obtained with proteins extracted from 3 independent infiltrations at 24°C or 4°C.


A
Reporter :
TaCBF4D-B9 PROM
GFP
NOS
Effectors :
CaMV 35S PROM
TaICE87
NOS
CaMV 35S PROM
TaICE41
NOS
B
PGII
1
Non-Infected
24°C
2
Non-Infected
24°C
3
PGII
Non-Infected
4
CBFp
24°C
5
CBFp
4°C
6
CBFp
CBFp + TaICE41
7
Non-Infected
CBFp + TaICE87
24°C
8
CBFp + TaICE87
24°C
9
CBFp + TaICE41
C
CBFp
CBFp + TaICE87
CBFp + TaICE41
24°C 4°C 24°C 4°C 24°C 4°C
GFP
RuBP

III-Figure 6: TaICE87 and TaICE41 enhance tolerance of *Arabidopsis* to freezing stress.

(A and B) Plants were grown for 3 weeks at 22°C (NA) or grown at 22°C then transferred at 4°C for 7 days (CA7). Wild-type plants (WT), empty vector transformed plants (PBI) and three lines over-expressing *TaICE87* (A) or *TaICE41* (B) were subjected to this experiment, and pictures were captured before the freezing test. The same plants were subjected to freezing and pictures were captured after a recovery period of 2 weeks (NA frozen to -6.5°C and CA7 frozen to -10.5°C).

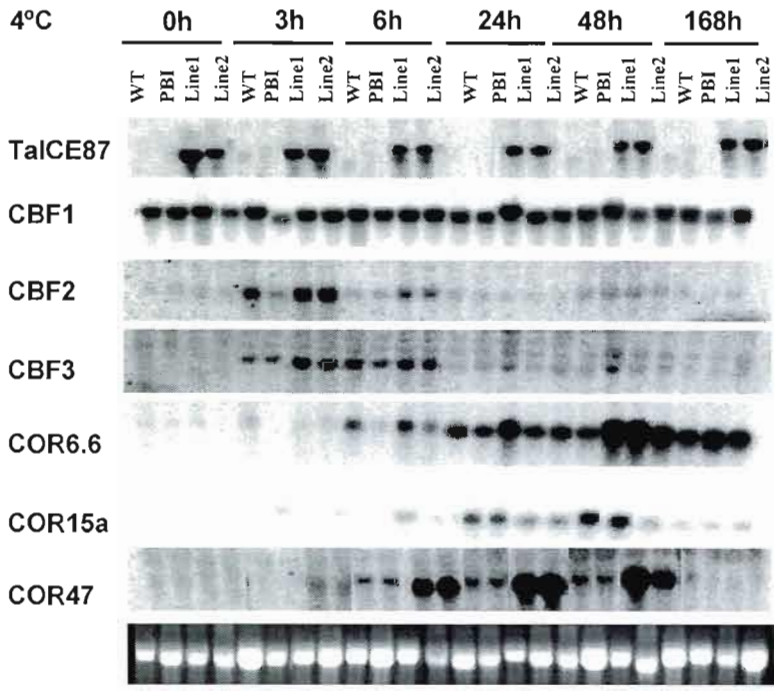
(C and D) Survival rate after freezing stress is expressed as a percent of surviving plants for *TaICE87* (C) or *TaICE41* (D) over-expressing lines. Statistical analysis was performed by one-way ANOVA, and the asterisks (*) indicate differences that are significant at the $P<0.001$ level.


A
NA
NA -6.5°C
CA7
CA7 -10.5°C
WT PBI L1 L2 L3
TaICE87
B
NA
NA -6.5°C
CA7
CA7 -10.5°C
WT PBI L1 L2 L3
TaICE41
C
Survival rate (%)
NA -6.5
CA -10.5
WT PBI L1 L2 L3
TaICE87
D
Survival rate (%)
NA -6.5
CA -10.5
WT PBI L1 L2 L3
TaICE41

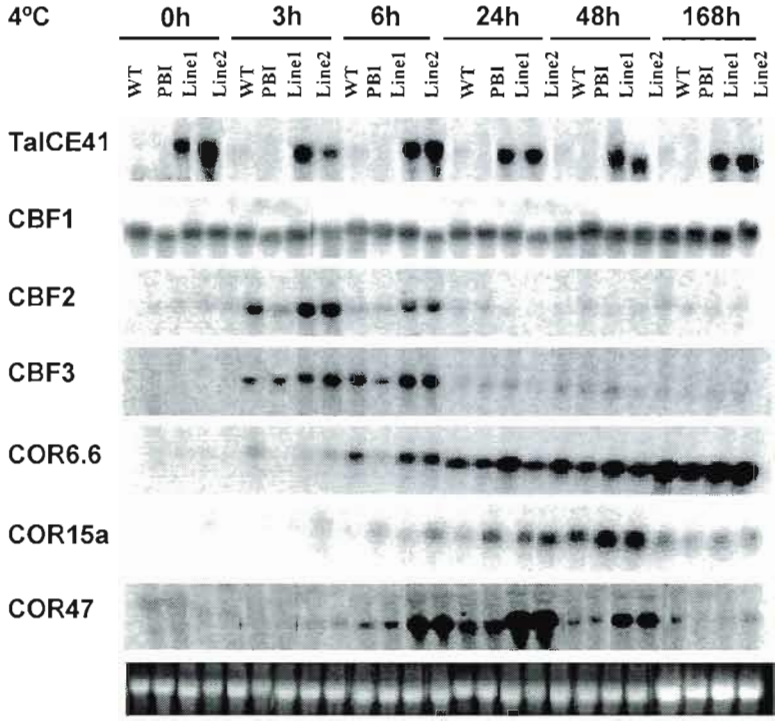
III-Figure 7: Effect of TaICE87 and TaICE41 over-expression on the accumulation of cold-regulated transcripts.

Total RNA was extracted from leaves of 15 day-old transgenic *Arabidopsis* plants grown under LD conditions, exposed at 4°C for 3 and 6 hours. Transcript levels were measured by northern blot. *TaICE87* overexpressing lines (panel A), *TaICE41* overexpressing lines (panel B). Two independent lines over-expressing *TaICE87* or *TaICE41* were used in this experiment. Each experiment was repeated three times using RNA prepared from two biological samples with similar results.

A


4°C
0h
3h
6h
24h
48h
168h
WT
PBI
Line1
Line2
TaICE87
CBF1
CBF2
CBF3
COR6.6
COR15a
COR47

B


4°C
0h
3h
6h
24h
48h
168h
WT
PBI
Line1
Line2
TaICE41
CBF1
CBF2
CBF3
COR6.6
COR15a
COR47

III- Supplemental DATA

III-Table S1: ICE-like genes and bHLH outgroup genes from different plant species used for phylogenetic analysis.

Organism	Name	Accession No.	Type of Sequences	Database
Arabidopsis thaliana	AtICE1	*AY195621*	mRNA	GenBank
	AtbHLH033	NM_101157	Genomic	GenBank
	AtbHLH061	NM_121095	Genomic	GenBank
Outgroup				
Brassica napus	*BnDY000939*	*DY000939*	EST contigs	GenBank
Capsella bursa-pastoris	*CbICE53*	*AY506804*	mRNA	GenBank
Citrus sinensis	*CsCN191080*	*CN191080*	EST	GenBank
Glycine max	*GmTC218613*	*TC218613*	EST contigs	TIGR
Gossypium raimondii	*GrCO111971*	*CO111971*	EST	GenBank
Hordeum vulgare	*HvBU983081*	*BU983081*	EST	GenBank
	HvlCE2	*DQ151536*	mRNA	mRNA GenBank
Lycopersicon esculentum	LeTC171424	TC171424	EST contigs	TIGR
Malus domestica	*MdEG631286*	EG631286	EST	GenBank
Oryza sativa	*Os11g0523700*	AK109915	Genomic	Genbank/RAP-DB
	Os01g0928000	*NM_001051807*	Genomic	Genbank/RAP-DB
Outgroup	*Os03g0135700*	AK103779	Genomic	Genbank/RAP-DB
Physcomitrella patens	*Pp2120038*	*estExt_fgenesh1_pg.C_2120038*	Genomic	JGI Physcomitrella patens subsp patens v1.1
	Pp510113	*estExt_Genewise1.C_510113*	Genomic	JGI Physcomitrella patens subsp patens v1.2
Pinus taeda	*PtaTC69556*	TC69556	EST contigs	TIGR
	PtaTC14490	TC14490	EST contigs	TIGR
Populus trichocarpa	*PtrLGXV000793*	C_LG_XV000793	Genomic	JGI Populus trichocarpa v1.1
	PtrLGXII1027	C_LG_XII1027	Genomic	JGI Populus trichocarpa v1.1
Outgroup	*PtrLGII0987*	*C_LG_II0987*	Genomic	JGI Populus trichocarpa v1.1
Saccharum officinarum	*SoTC59580*	TC59580	EST contigs	TIGR
	SoTC51628	*TC51628*	EST contigs	TIGR
Solanum tuberosum	*StDV627678*	*DV627678*	EST	GenBank
Triticum aestivum	*TaCA501920*	*CA501920*	EST	GenBank
	TaCK208335 *TaICE41* *TaICE87*	*CK208335*	EST	GenBank/FGAS
Zea mays	*ZmTC348661*	*TC348661*	EST contigs	TIGR
	ZmDV024434	*DV024434*	EST	GenBank

III-Table S2: Sequences of ICE1, ICE1-like proteins and outgroup bHLH genes used for phylogenetic analysis.

>AtbHLH033
MNSDGVWLDGSGESPEVNNGEAASWVRNPDEDWFNNPPPPQHTNQNDFRFNGGFPLNPSENLLLLLQQ
SIDSSSSSSPLLHPFTLDAASQQQQQQQQQQEQSFLATKACIVSLLNVPTINNNTFDDFGFDSGFLGQ
QFHGNHQSPNSMNFTGLNHSVPDFLPAPENSSGSCGLSPLFSNRAKVLKPLQVMASSGSQPTLFQKRA
AMRQSSSSKMCNSESSSEMRKSSYEREIDDTSTGIIDISGLNYESDDHNTNNNKGKKKGMPAKNLMAE
RRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIDYLKELLQRINDLHTELESTPPSSSSLHPLTPTP
QTLSYRVKEELCPSSSLPSPKGQQPRVEVRLREGKAVNIHMFCGRRPGLLLSTMRALDNLGLDVQQAV
ISCFNGFALDVFRAEQCQEDHDVLPEQIKAVLLDTAGYAGLV
>AtbHLH061 (outgroup)
METELTQLRKQESNNLNGVNGGFMAIDQFVPNDWNFDYLCFNNLLQEDDNIDHPSSSSLMNLISQPPP
LLHQPPQPSSPLYDSPPLSSAFDYPFLEDIIHSSYSPPPLILPASQENTNNYSPLMEESKSFISIGET
NKKRSNKKLEGQPSKNLMAERRRRKRLNDRLSLLRSIVPKITKMDRTSILGDAIDYMKELLDKINKLQ
EDEQELGSNSHLSTLITNESMVRNSLKFEVDQREVNTHIDICCPTKPGLVVSTVSTLETLGLEIEQCV
ISCFSDFSLQASCFEVGEQRYMVTSEATKQALIRNAGYGGRCL
>AtICE1
MGLDGNNGGGVWLNGGGGEREENEEGSWGRNQEDGSSQFKPMLEGDWFSSNQPHPQDLQMLQNQPDFR
YFGGFPFNPNDNLLLQHSIDSSSSCSPSQAFSLDPSQQNQFLSTNNNKGCLLNVPSSANPFDNAFEFG
SESGFLNQIHAPISMGFGSLTQLGNRDLSSVPDFLSARSLLAPESNNNNTMLCGGFTAPLELEGFGSP
ANGGFVGNRAKVLKPLEVLASSGAQPTLFQKRAAMRQSSGSKMGNSESSGMRRFSDDGDMDETGIEVS
GLNYESDEINESGKAAESVQIGGGGKGKKKGMPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRAS
ILGDAIDYLKELLQRINDLHNELESTPPGSLPPTSSSFHPLTPTPQTLSCRVKEELCPSSLPSPKGQQ
ARVEVRLREGRAVNIHMFCGRRPGLLLATMKALDNLGLDVQQAVISCFNGFALDVFRAEQCQEGQEIL
PDQIKAVLFDTAGYAGMI
>BnDY000939
ISDEHNKGKKKGMPAKNLMVERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIDYLKELLQRINDL
HTELESTPSSSSGLTPSPQTLPYRVKEELCPSSSSLSSPKGEQARIEVKLREGKVVNIHMFCGRRPGL
LLSTMRALDDLGLDVQQAVISCFNGFALDVFRAEQCQEGHEVFPEQIKAVLLDTAGYSGLL
>CbICE53
MVLDGNNGGAVWLGGGGEREENEEGSWGRNQEDGSQFKPMLEGDWFSNPPHPQDLQMLQNHQDCRFLG
GFPFNPNDNLLLQHSIDSSSSCSPSQAFSLDPSQQNQFLSTNNKSCLINVPSSANPFDNAFEFGSESG
FLSQIHAPMSMGFGSLTQLGNRDLSSVPDFLSARSLLAQDHNSSNSVLCGGGGGFTAPLELEGFGSPA
NGGFVGNRAKVLKPLEVLASSGAQPTLFQKRAAMRQSSRSKMGNSESSGMRRLSDDGDMDETGIGVSG
LNYESDEPNESGKAAESVQIGGGKGKKKGMPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASIL
GDAIDYLKELLQRINDLHNELESTPPGSLPPTSSSFHPLTPTPQTLSCRVKEELCPSSLPSPKGQQAR
VEVRLREGRAVNIHMFCGRRPGLLLATMKALDNLGLDVQQAVISCFNGFALDVFRAEQCQEGQEILPD
QIKAVLFDTAGYAGMI
>CsCN191080
KGKKKGLPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKELLQRINDLHNELES
TPPGSALTPSTSFYPLTPTPPALHSRIKDELCPSSLPSPNGQPARVEVRVREGRAVNIHMFCSRRPGL
LLSTMRALDNLGLDIQQAVISCFNGFAMDVFRAEQCKEGQDVHPEQIKAVLLDSAGFHGMM
>GmTC218613
FSPRGKRKEKEKDEEVEDLSFDGSGLNYDSDDLTESNYNNVSEGNTGKNGGVSSNANSTVITGLDQKG
KKKGMPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKELLQRINDLHNELESTP
VGSSLTPVSSFHPLTPTPPTLPSRIKEELCPSSLPSPNGQPARVEVRLREGRAVNIHMFCARKPSLLL
STMRALDNLGLDIQQAVISCFNGFAMDIFRAEQCKEGQDVHPEQIKAVLLDSAGYNGM
>GrCO111971
KGKRKGLPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKELLQRINDLHNELES
NPGSSSLTPTTSFHPLTPTPATLPCRIKDELCPSSLPSPNGQPARVEVRLREGKAVNIHMFCGRRPGL
LLSTMRALDSLGLDIQQAVISCFNGFAMDIFRAEQCKEGQDIHPEQIKAVLLDSAGFNNMI

>HvBU983081
MDRASILGDAIEYLKELLHKISDLQNELESSPSMPSLPSTPTSFHPLTPTLPSLPSRVKEELCPSALP
SPTGQQPTVEVRLREGQAVNIHMLCPRRPGLVLSAMKAIESLGLDVQQAVISCFNGFALDVFKAEQCK
DGPGLQPEEIKAVLLQSAGFHPTM
>HvICE2
AFGDGMGWDDEDELDQQSMDASSLGVSASLENAAVGAPGGGGGGGNGKGKKKGMPAKNLMAERRRRKK
LNDRLYMLRSVVPKISKMDRASILGDAIDYLKELLQRISDLHSELESAPSSAALGGPSTANTFLPSTP
TLQPFPGRIKEERCPPAPFPSPSGQQATVEVRMREGQAVNIHMFCARRPGILLSTMRALDSLGLDIEQ
AVISCFDGFAMDVFRAEQCREGPGLLPEEIKAVLLHCAGLQNAM
>LeTC171424
GKRKWGSGEELDDVSFDGCTLSYDSDDLTENVTNKVDDTVKNGGNSSNATSTVTCGNQKGKKKGLPAK
NLMAERRRRKKLNDRLYMLRSVVPRISKMDRASILGDAIEYLKELLQKINDLHNELESTPPSSSLTQT
TSFYPLTPTGPALPGRIKEELYPSSFASPLSSPTGQPARVEVKAREGRAVNIHMFCSRRPGLLLSTMR
ALDNLGLDIQQAVISCFNGFALDIFRAEQCKEGQDFHPDQIKAVLLDSAGCHGMI
>MdEG631286
KKGLPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKELLQRINNLHNELESIPP
GSALTPTGNTFHPLTPTPATLPNCIKEELCPSSLPSPNGQAARVEVRLREGRAVNIHMFCGRRPSLLL
STMRTLDNLGLDIQQAVISCFNGFAMDVFRAEQCKEGQDVHPDQIKAVLLDSIGFHGMM
>Os01g0928000
MDEAEAAAAAAAKMDELAGGGGGGGGDWSYLAADALAAASFTAFPFHHHHHHHHRDVLSASTPSSLLLN
MDAATAAAMFDFQAAFPSSSVPPPPPTTTAALPPFHDFASSNPFDDAPPPFLAPPGQKLGFLGPPGGA
FGGGMGWDDDDEIEQSVDASSMGVSASLENAAPVAAGGGGGGGGGGGRGKKKGMPAKNLMAERRRRKK
LNDRLYMLRSVVPKISKMDRASILGDAIEYLKELLQRINDLHNELESAPSSSLTGPSSASFHPSTPTL
QTFPGRVKEELCPTSFPSPSGQQATVEVRMREGHAVNIHMFCARRPGILMSTLRALDSLGLGIEQAVI
SCFNGFAMDVFRAEQCRDGPGLGPEEIKTVLLHSAGLQNAM
>Os03g0135700
MELDEESFLDELMSLRRDGSAPWQAPPYPGGGGGGGGGGMMMSDLLFYGGDGGSAEARGGMDASPFQE
LASMAAPPPQHPHEEFNFDCLSEVCNPYRSCGAQLVPSEAASQTQTQLTPLRDAMVAEEETSGDKALL
HGGGGSSSPTFMFGGGAGESSEMMAGIRGVGGGVHPRSKLHGTPSKNLMAERRRRKRLNDRLSMLRSI
VPKISKMDRTSILGDTIDYVKELTERIKTLEEEIGVTPEELDLLNTMKDSSSGNNNEMLVRNSTKFDV
ENRGSGNTRIEICCPANPGVLLSTVSALEVLGLEIEQCVVSCFSDFGMQASCLQEDGKRQVVSTDEIK
QTLFRSAGYGGRCL
>Os11g0523700 (outgroup)
MLPRFHGAMWMQDDGGGDQEHGQAAPPGQEQHHHDQHLMALAAAAAGGAGFGAAQAPAPLLDEDWYFD
AAGGGGGGAHGSMMLGLSSVHGGIGAGTSGGGHGQQFSLLNMGAAAAPFDVSGFDLGIACGGVGGGGD
VVSFLGGGNASNTALLPVGNAGFLGTFGGFGTAASQMPEFGGLAGFDMFDAGAVNTGGSSSSSSAAAA
AASASAHVSNTAPFSGRGKAAVLRPLDIVPPVGAQPTLFQKRALRRNAGEDDDDKKRKAAAGAGAGAL
SADGADMVLDDGDDDGLSIDASGGLNYDSEDARGGEDSGAKKESNANSTVTGDGKGKKKGMPAKNLMA
ERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKELLQKINDLQNELESSPATSSLPPTPTSF
HPLTPTLPTLPSRIKEEICPSALPSPTGQQPRVEVRLREGRAVNIHMFCARRPGLLLSAMRAVEGLGL
DVQQAVISCFNGFTLDIFKAEQCKDGPGLLPEEIKAVLMQSAGFHTMI
>Pp2120038
MRAASFCKGVLDEEWYTPETSLMELSSCSSPYGAQDARSNFSLLDSSLNYDNDKLMANFRPAPSGTLG
IGQPERNCILSDLACTGQSSSVGLVSTHSFPVLSRSNAGAFYASENEGGKFVRSVLVLLLVMLLRLPL
GITAPWLKLVGKPEPRMSIVVLRRRCLPAKSQHVASYLDAMDRASILGDAIEYLKELLQRINDIHNEL
EEAKLEQSRSMPSSPTPRSTHQGYPTAVKEECPVLPNPESQPPRMEVRKREGQALNIHMFCARRRGLL
LSTVRALDALDLDVQQAVISCFNGFALDLFRAEAHLVFGRRTLTAAIKAPCDPGVRIKMVSQCLEFWN
RITSRVIFESSMEQSSNRNRRHAAAVAASIHEPR
>Pp510113
MSSFRPAPSALSMGLESNRSLEDLVCTGQGSSNVGLLSSLSPGGQLGRSTVMESFSSGLPTSFNQGII
NAGGSITNITSSNINNVRSNFPLMASPSNFSDAYRARSVSEDKSGKVVGSGGPRNELVPYHRNKGAET
RSHGQGQQTLFLKRAASRRCAGSSGTVSPVSKSPPRVVTSASNDSSVDTPDKDSPHPRNAHLQSASGR

LNINSGSDDPNDMGLDGDDYDAKDDDDLDESGDGSGGPYEVEEGAGNGADQSIGKGNGKGKRGLPAKN
LMAERRRRKKLNDRLYTLRSVVPKITKMDRASILGDAIEYLKELLQRINEIHNELEAAKLEQSRSMPS
SPTPRSTQGYPATVKEECPVLPNPESQPPRVEVRKREGQALNIHMFCARRPGLLLSTVKALDALGLDV
QQAVISCFNGFALDLFRAEAKDVDVGPEEIKAVLLLTAGCDLHSLQ
>PtaTC14490
RGRKNGLPAKNLMAERRRRKKLNDRLFMLRSVVPKVSKMDRASILGDAVEYLKELLQRINDLHIELMA
GSSNSKPLVPTMPDFPYRMNQESQASLLNPEVEPATVEVSTREGKALNIHMFCSKKPGLLLSTLRALD
ELGLDVKQAIISCLNGFALDVFRAEQSMGGDVTAEEIKALLLHTADNEDGL
>PtaTC69556
RAEFGTRAVNNNSNGGGDKGKKKGLPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIE
YLKELLQKINDLHNELESTSQGPVLPGTSNFHPLTPTTPSLPCRVKEECPTSLPSPNAQPARVEVRMR
EGHALNIHMCCARRPGLLLSTMRALDGLGLDVQQAVISCFNGFALDVFRAEQAKEGEIAPEEIKAVLL
HTASCHTAI
>PtrLGII0987 (outgroup)
MEINGHDFLEELIALSRESWQPTPNYPSEMNELFSGSFNHGCFEEIPATLPQTSFCPEGLISSPLKQD
FNNYYFNEVFCPFGDEFSAPQFTDEFSSAPQFTDSSYNTLDTPPFPVQDDTPMSMMEDEELGLLANDQ
QNLQMQGTCKVEPIQSPEVSAFNAGICPERKIRGKKMEGQPSKNLMAERRRRKRLNDRLSMLRSIVPK
ISKMDRTSILGDTIDYMKELLERINSLQQEIEVGSEELKMISIFKDTKPNEIVVRNSPKFEVERRNED
TRIDICCATKPGLLLSSVTTLETLGLEIQ
>PtrLGXII1027
MLYRLNSNNIWLEDHKEEQDSTTNHHHHNNITNTAAGCGGVMLEGKEEMGSLSTFKSMFEVEDEWYVT
NNNSTIHQNHQDSIKDLTFSPSLVDPDNLLLHQVDSSSSCSPSSSVFNNLDPSQVHYFMHPKPTLSSL
LNVVSNNPLEHGFDLSEIGFLENQGTNSTTTANVSSLLNRGSGVLGNLGNFTDLSSNSQISIPNLCSD
PQFSSSRMLQLPENGPGFNGFRGLDEISGNQLFFNRSKLLRPLETYPSMGAQPTLFQKRAALRKNLGE
VERDKGKREMTQISEEKDKKRKFSSGDDFLEDVSFDGSGLNYDSDEFTENTNLEETGKNGGNSSKANS
GVTGGGVDQKGKKRGLPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIDYLKELLQRI
NDLHNELESTPPSSSLTPTTSFHPLTPTPSALPSRIMDKLCPGSLPSPNGQPARVEVRVREGRAVNIH
MFCGRKPGLLLSTMRALDNLGLDIQQAVISCFNGFAMDIFRAEQCKEGQDMHPDQIKAVLLDSAGFHG
AM
>PtrLGXV000793
MGSLSTFKSMLEVEDEWYVSNNNNTIHQTHQDSIKDLTFSPGLGDPDNLLLHQVDSSSSCSPSSSVFN
NLDPSQVHYFMHPKPSLSSLLNVVSNNPLEHGFDLSEIGYLENQGTNSAATANISIPNLCSDPQFSSS
RMLQLPENGPGLTSFRGFDENSGNQLFLNRSKLLRPLETYPSMGAQPTLFQKRAALRKNLENDKKRKF
SSGDDFLEDVSIDGSGLNYDSDEFTENTKVEEIGKNGGISSKANSGVTGGVDQKGKKKGLPAKNLMAE
RRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKELLQRINDLHNELESTPPSSSLTPTTSFHP
LTPTPSALPSRIMDKLCPSSLPSPNSQPARVEVRVREGRAVNIHMFCGRKPGLLLSTMRALDNLGLDI
QQAVISCFNGFAMDIFRAEQCKEGQDMHPDQIKAVLLDSAGFHGMM
>SoTC51628
NPFDDAGHFLGGPPPLPPPPAAQQQGQKGGFFAPPPGSDFNDTGMSWDDEDEIDQSVDTSSMAISAYM
ENAAGAAAGGSGAGCGSGRGKKKGMPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIE
YLKELLQRISDLHNELESASSSSFVGPTSASFNPSTPTLQTFPGQVKEELCPGSFPSPTGQQATVEVR
MRDGHAVNIHMFCARRPGILLSTMTALDSLGLDIEQAVISCFNGFAMDVFRAEQCADGPGMVPEEIKA
VLMHTAGLHNAM
>SoTC59580
MPDFGGLGGFDMFNNGAGSSSAAPPPASVSLTAPFSGRGKAAVLRPLEIFPPVGAQPTLFQKRALRRN
AGEEDDDKKRKAEAITAAAGASSAGGDDTVLDDADDDDGGSIDASGLNYDSEDARGVEDSEKKDGKDS
NANSTVTGGGTGDGKGKRKGLPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKE
LLQKINDLQNELESSPSTASLPPTPTSFHPLTPTLPTLPSRVKEELCPSALPSPTSQQPRVEVRMREG
RAVNIHMLCTRRPGLLLSAMRAIEGLGLDVQQAVISCFNGFSLDIFKAELCNEGPGLLPEEIKSVLLQ
SAGFHGVMP
>StDV627678

KGKKKGLPAKNLMAERRRRKKLNDRLYMLRSVVPRISKMDRASILGDAIEYLKELLQKINDLHNELES
TPPSSSLTQTTSFYPLTPTGPALPGRIKEELYPSSFASPLSSPTGQPARVEVKARDGRAVNIHMFCSR
RPGLLLSTMRALDNLELDIQQAVISCFNGFALDIFRAEQCKEGQDFHPDQIKAVLLDSAGCHGMI
>TaCA501920
PAGTGPEFPGRPTRPGKKKGMPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKE
LLHKISDLQNELESSPSMPSLPPTPTSFHPLTPTLPALPSRVKEELCPSALPSPTGQQPTVEVRLREG
RAVNIHMLCPRRPGLVLSAMKAIEALGLDVQQAVISCFNGFALDVFKAEQCKDGPGLQPEEIKAVLLQ
SAGFHPAM
>TaCK208335
TGKKKGIPPKNLMAERARRKKLNARLYAFRSVVPRISKMDRASILGDAIEYLKELTQKINVLQNELEA
SPSTSSLPPTPTSFRPPTPTMPALPSRVKEELTSSPAQEPRVEVKLREGRVVNIRMMCSRRPGVVHSS
LKALEGLGLDVQQAVISYFNDFTLDVSKAECKDGPGPQAEEIKAVLLQSAGFHPAA
>TaICE41
MENSAAAVGVEKEDELVGGGGGDWGYLTSEAMATAGFPAFGFPCGARGGVTPAPTSASLLMSMEHAAL
FDYNAAFPSSSSSAAAAPPAYHDFGSGGNPFNVDAPPFLLEAPPPLTVAPGGQKGGFLAPPLSAFGDG
MGWDDEDELDQQSVDASSLGVSASLENAVVGAPGGGGGGGGGNGKGKKKGMPAKNLMAERRRRKKLND
RLYMLRSVVPKISKMDRASILGDAIDYLKELLQRISDLHSELESAPSSAALGGPSTANSFLPSTPTLQ
PFPGRIKEERCPPAPFPSPSGQQATVEVRMREGQAVNIHMFCARRPGILLSTMRALDSLGLDIEQAVI
SCFDGFAMDVFRAEQCREGPGLLPEEIKAVLLHCAGLQNAM
>TaICE87
MLDDDSWYFNPGAVGDAAGNGSMIPAPATMEASGSSSGFGDASQMFPLLNPGVGGTGPLDVPGFDLDI
SGDLSAFLGAGNAPNTSLLPRGNTDFLGSFGGFGTAPAQTTDFGGLAGFDLFDTGAQGWSGPSSEGPA
APAPQTAPFSGQGNAKAMRPLDTFPASGAQPTLFQKRALRRNAGEEDGGRKRKAAEPDIILDDADDDI
ISIDASGLIYDSEDGRGVEESGRKDGNESNANSTVTGGATAEGNAKKKGMPAKNLMAERRRRKKLNDR
LYALRSVVPRISKMDRASILGDAIEYLKELKQKINVLQNELEASPSASSLPPTPTSFHPLTPTTPTMP
ALPSRVKEELASSAAQEPCVEVKLREGRVVNIRMMCSRRPGVVHSSLKALEGLGLDVQQAVISYFNDF
TLDVFKAEQCKDGPGPQPEEIKAVLLHCAGFHPAV
>ZmDV024434
GASSGGGGDTVLDDADDDDGGSIDASGLNYDSEDARGVEDSGKKDGKDSNANSTVTGGATGDGKGKRK
GLPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKELLQKINDLQNDLESSPSTA
SLPPTPTSFHPLTPTLPTLPSRVKEELCPSALPSPTSQQPRVEVRMREGRAVNIHMLCARRPGLLLSA
MRAIEGLGLDVQQAVISCFNGFSLYIFEAELCKEGPGLLPEEIKPVLLQSAGF
>ZmTC348661
GAPPPVPAAAAPQQQGQKGGFFAPLPASDFNDAGMSWDDEDEIDQSVDASSMAISASMENAAGAVAGA
SGAGGGSGRGKKKGMPAKNLMAERRRRKKLNDRLYMLRSVVPKISKMDRASILGDAIEYLKELLQRIS
DLHNELESAPSSSLVGPTSASFNPSTPTLQTFPGQVKEELCPGSFPSPTGQQATVEVRMREGHAVNIH
MFCARRPGILLSTMTALDSLGLDIEQAVISCFNGFAMDVFRAECADGPGMVPEEIKAVLMHTAGLHNA
M

III-Table S3: Number of potential MYC variant binding sites in the wheat and Arabidopsis CBF promoter sequences.

The four possible nucleotides on the 3rd base are named MYC 1 through MYC4 and the 4th base is used to name the MYC variant. Ta promoters' nomenclature is according to the phylogenetic analysis from Badawi et al. (2007). Light-blue: *Arabidopsis* cold-regulated CBFs. Dark-blue *TaCBF* promoter used for different experiments in this publication. Orange: Myc2a and Myc4g in the different promoters.

Promoter	MYC variant															
CANNTG MOTIF (3rd base)	MYC1 A				MYC2 C				MYC3 G				MYC4 T			
CANNTG MOTIF (4th base)	A	C	G	T	A	C	G	T	A	C	G	T	A	C	G	T
AtCBF1							2	1					1			
AtCBF2					1		1	1								1
AtCBF3		1		1	1	1							1			2
AtCBF4			1				1									
AtDDF1	1	1	1	1					3					1		2
AtDDF2	2			1									1			
TmCBF5	1				1		3			1		1				
TaCBFIIIa-6.1			1	1			1	1								
TmCBF18						1	2			1		1	1			
TaCBFIIIc-3.2										1						
TmCBF3								1		1						
TaCBFIIIc-B10										1						
TmCBF10		1		1	1		2		2	1		1	1	1		1
TmCBF13	1			1	1			1	2	1						
TmCBF12	2	1			1	1		1	1		1					
TmCBF15		1				1		3	1							2
TmCBF16			2			1		1					1			1
TmCBF17									1				1			1
TaCBFIIId-A19					1	1		2		1				4		
TmCBF2			1	1					1		1	1				
TaCBFIVb-D20		3						2	1						1	
TmCBF14															1	1
TaCBFIVd-4.1			1						1		1	1			1	
TaCBFIVd-B9	2		2	1	1			1			1			3	2	
TaCBFIVd-D22			1												1	

III-Table S4: Primers used in different experiments.

A) Primers used to generate constructs for the production of recombinant TaICE87 and TaICE41 proteins. The restriction enzyme cutting sites used for cloning in pDEST15 are underlined.

TaICE87-F	*5'-GAGATATCATGCTTGATGACGACAGCTGGTA-3'*
TaICE87-R	*5'-GAGAATTCCTAGACCGCTGGATGGAACC-3'*
TaICE41-F	5'-TGACCATGGAGAACTCGGCGGCGG-3'
TaICE41-R	*5'-GAGAATTCCCTACATCGCGTTCTGGAGACC-3'*

B) Primers used to generate constructs used in the transient expression assays. The restriction enzyme cutting sites used for cloning in pGreenII0029 are underlined.

a. Reporter construct

TaCBFIVd-B9	*5'-TTAAGCTTCCAGCGCCACAGCATTCGGCGT-3'*
TaCBFIVd-B9	*5'-TCTCTAGACGGAGATGCTCGTGAGAGGTGTG-3'*

b. Effector constructs

TaICE87f1	*5'-TAGGATCCATGCTTGATGACGACAGCTGGTA-3'*
TaICE87r1	*5'-GAGAATTCCTAGACCGCTGGATGGAACC-3'*
Ta ICE41f1	*5'-TAGGATCCATGGAGAACTCGGCGGCGG-3'*
Ta ICE41R1	*5'-GAGAATTCCCTACATCGCGTTCTGGAGACC-3'*

C) Primers used to generate the *Pro$_{35S}$:TaICE87 and Pro$_{35S}$:TaICE41* constructs for *Arabidopsis* transformation. The restriction enzyme cutting sites used for cloning in pBIN19 mGFP-ER are underlined.

TaICE87f1	*5'-TAGGATCCATGCTTGATGACGACAGCTGGTA-3'*
TaICE87r1	*5'-GAGAATTCCTAGACCGCTGGATGGAACC-3'*
Ta ICE41f1	*5'-TAGGATCCATGGAGAACTCGGCGGCGG-3'*
Ta ICE41R1	*5'-GAGAATTCCCTACATCGCGTTCTGGAGACC-3'*

D) Primers used for the generation of probes for the different *CBFs* and *CORs* genes.

AtCBF2F	*5'-ACCTTGGTGGAGGCTATTTATACG-3'*
AtCBF2R	*5'-CATTTGCATTTGACAACAACTTTTACC-3'*
AtCBF3F	*5'-CAGAGCGAAAATGCGTTTTATATGCA-3'*
AtCBF3R	*5'-TAATTTACACTCGTTTCTCAGTTTTACA-3'*
AtCOR6.6F	*5'-GTGTTAACTTCGTGAAGGACAAG-3'*
AtCOR6.6R	*5'-CAAACGTAGTACATCTAAAGGGAGA-3'*
AtCOR15aF	*5'-GATACATTGGGTAAAGAAGCTGAGA-3'*
AtCOR15aR	*5'-CGGTGACTGTGGATACCATATCTT-3'*

III-Table S5: Nucleotide sequences of CBF promoters from *Arabidopsis*, *Triticum monococcum* and *Triticum aestivum.*

>AtCBF1, AAC99369, 1442 bp, 4 MYC.

AACATATCATCA**CACGTG**GAATGAGAACGAGTTTCGACTTTTCAAATATGCCATAAAGCCTCAATTAT
CTTCTTATCTAGCTTGAATATGCAACAAAAAGCTATTAAGATATTCATAAAATAGAGGCGTCTCAAAT
CTACCAACAAAAAGCTACAAAAGATCCAGTCCAATCCACTGAAGAATCCCAAAACAGAGTAGAAACCC
AAACAACCCGATTCAGCAAATAATTCAAAAACGAACGTCCGTACGATTTTCCAAAAACAGAAAGATGG
GTTCCACAAGATAATGCGTGGGGACGTCAAAATCCTCTAAACCCGTGGTTCGGCCGCGGAAACACTGT
CCCTACCTTCCCCACCAGCCTTCACTGGCCCCATACGTCACTCTCAAGCTTTACTTTCTATTTTCCAC
TAAAGCCAATTTTGTTGTTTTCTTCACCTTACCACTCTTTTTTTCCCTCTTTGTTGTGTCTTCTTTTC
TCCTAAATGTCAATAACGTGAGAGCGAGAGGTAACGAGAGAGATATTTTTGTCAGCGAATATATTTCA
TGCATATCTTATTGTGAAGATTTTTTATACCTTTTTTTTGTCAATACAATATAGCTATTATTGAGATT
GAGATATTTTGTGGAATTATTGGGATTCAAGATAACTTGCTATTTTGTATTGGTCTTATCCTTCGCTT
AGTCCTGTCCTGGTCCATTTACATGTTTTTGGTTATAGTTTGTTTAAACTGAATAATTTTGTTCAT**CA**
TATGCATTAATGACTCATTTTTAACCGTCCATCGAAATTGATAATTATCCATTACCAAATCTGATTAA
TTTTTTTAAAAAATCAAGCTTTTCTATATTGTAGTATTATTTTTGGTTAAATATTAGGACATCTACTT
CCAATACAAATACTACATGAGTATTTAAAATATCATTTCACAGAGATATTTATGTCTATTATGTTATA
GACGGGTGACAATTAATGACAATTTGTTTATTCATAGGAATTTAAAAACGATTGTAACAACAGCAGCC
AGCCAACCACACAGGCACACACTCGATAGAATTTAAAGAACTCATAAAGGTTAACGAGTGAAGAGTCA
AAAGTCTCTTTACAAGGGTCAAAGGACACACGTCAGACAGCGAGTGGAACATCGTGGGATTGCTTCGC
TATGTACTATA**CACGTG**TCATTCACAGAGACAAAAACTCCGTGTGCACCCCACATATCCGTTATCTCT
CCTCCGGCCAATATAAACACCAATTCTCACTCTCACTTTTTATACTAACTACA**CACTTG**AAAAAGAAT
CTACCTGAAAAGAAAAAAAAGAGAGAGAGATATAAATAGCTTTACCAAGACAGATATACTATCTTTTA
TTAATCCAAAAAGACTGAGAACTCTAGTAACTACGTACTACTTAAACCTTATCCAGTTTCTTGAAACA
GAGTACTCTGATCA

>AtCBF2, AAC99371, 1500 bp, 4 MYC.

ATAAGCGGGGTTAATAGATCAACCACAATCAATTTAATTTGGACTTTAGAATTAATAAAATTGTTTAC
TTCGTAATTATTATTATTTTTGTTGTTCTGGCAAATCTGATAATCCAGATTATTATTAGACAAGTAGC
AAAGGGACGGTGAACATTTATGATTTTAATTTGTATGTTGTGAGGAAAACAAAACAAATAAGTTCTGT
AAAAAAGGTTTACCTTTCTACTTTGCCGGAAAACTCAACTCACGGTGGCGTCCGGCGAGTTTTCAGAC
CAAAAAGAAGGTTGGAAGAAATGAAGATGAAGAGGAGAGGACAAAAGATAGAGATGGTGGTTGAACAA
AAGAAGAGTAAAGAGGACGAAGACGCTCTAAGTCTAAGCCAAGGGGGAGAAGAAGAGAAGAGGTATGA
GGAGGAACCATACTTTTGTTAGAGAGATGCTGGAAATTGTGATCAACTACATGCAAAATGTCTTTTCG
CCTAACCACTTACCATATTTGATATTTTCCTTTTGCCAAATTACACAAACCCTATCTTGTCTCTCACA
TATATATCCAATTAATACACCCCTGC**CACTTG**TTAATTCTCGACCATGTATGTATACTTATGTAAAGA
ATATCCAAAAGCTTTCTTTTTGTTCCTTCGATTTTAAGCAACTTGTGTTCTCATTTCTCAATATATTA
AAGAAATCCTGAGTAAAAGTTATAGCCTCCGTGAATCTTAGGAAATTACTCTAGCATATTCAAATTTT
TTGAGACAATATATAAATTTTTCTGAATAATTAAATTTACATATCTATGCTACGAAACTTGATTAATT
AAATTAAATATATATATATAATAATAATAATAATAATAATATAACATTTTTTTTAGGACACAAATATC
TAATCTCACTATACTCTAGAAGTATTTGCAATGCACGATATGTGAATGGAGAAAAGACAGAAAGAG**CA**
TTTGAAAATATCTCGTTTCACGGATCATTATGTCTAATTATTTTACCATAGAAAAGCGACAATTATAA
ACAATTTGTTATTCGTGGAAAAATAATATTTAATAATGGTTGTCGTACCCTATAAACTACAGCCACAC
ATTCATACAATAAGAAGTTAAAAAAATTCATACCCTAAAGGCATCAACCAGTGAAGGGTCAGAAACTT
CCCAAGATGGGTCAAAGGA**CACATG**TCAGATTCTCAGTGATTGACAGCCTTGATAATTACAAAACCGT
GGGATCGCTTAGCTGTTTCTTATC**CACGTG**GCATTCACAGAGACAGAAACTCCGCGTTCGACCCCACA
AATATCCAAATATCTTCCGGCCAATATAAACAGCAAGCTCTCACTCCAACATTTCTATAACTTCAAAC
ACTTACCTGAATTAGAAAAGAAAGATAGAGAGAGAAATAAATATTTTATCATACCATACAAAAAAAGA
CAGAGATCTTTCTACTTACTCTACTCTCATAAACCTTATCCAGTTTCTTGAAACAGAGTACTCTTCTG
ATCA

>AtCBF3, AAC99370, 1500 bp, 7 MYC.

TAAGTATGTGATCAAAAGAAAGTATGTAATCAAAAGGGTTAGCACGAGTACCTTGGGAGGAAATTCTT
CTAATTATGAATTATGCAAGAATTTTCGTCAAGGGAAGGTGGGGAAGAGGTAGCTAAATTAAAGAATA
GAGAAT**CATATG**ACTAAGGACGTGGTGGTTGAAGGAAATGAGAGAATACATGAAGAAGAGAAACTTCT
TTGAGTGAGAAGGAAGTGCGCTGGCTGAAGGCAATAGAGAGAAAAGAGTTTCGAGTGAGAGAGAGGGC
GTTGAGATTGTGATCAACTTAATGTAATATGTTCTTTTATTACATTTTCTTTTTGTCATATACTCAAA
CCTTTTACTATTTTGTCTCATAAATCTAACACACCCCAC**CATTTG**TTAATGCATGATGGTAGAAAATA
TTAAATATAATTAACTACTTTTATGTGATCAAAATTAGGTTTCAGACTCGTTTCGCGATCCGATTTAC
AATTA**CAACTG**CATGCTTCTAATTGATCTAAATTCTAAGTTTTTTATACATATTAAAAAAATAACTTT
TTGTTAAATTCTCAATCATCATTTTTGTGATTAACAATTTTTTATAACTCTAAACCAATAATATTTGA
TTATTTATTTTATATGTATAATGATGATTGAGAATTTTAATTAGCAGTCTATTTAGGGTTTTCCTAAA
GTTACAATATGTTGTTACCCTTCTAGTTAAATTTTCCAAAATACCATATTTCATAACTTTTCAAACTG
TTTATTAATTCAACCGTAAAAAGCACTAAAATGTTA**CATTTG**ATCATTCACCCAAATTAAATTCAAAA
GTTTTTCCGCCAAAACTACTTGGTGACTTACGTGCTTATATACGGACGACTATTATTATGTTCTATAC
TTTTTTATACTTTGTTGCACAAATATCTACTCTCCCAATTCATATTCTAGAAGGATGTGCTATAAGAA
TGGGAGAAATTACACAAGAAGAGCATCTTTAAATATCCTCTCACAATCTTTATGTCTAATACACGGGT
GAACAATTAACGACAATTTCTTTATTCAGGAATATAATAATGAATAACGGTTACCCTACACCTAGTAC
ACTAAATCCTTAACAGCCACACATTCATACGCAAAGAGTTTATAAAACTCATAAAGGTATAATAATAA
CGAGTGAATAAGTCAAAAAAAAGTCTTCTCTGGA**CACATG**GCAGATCTTAATGAGTGAATCCTTAAACT
ACTCATTTTA**CAATTG**CTTCGCTGTGTATAGTTTACGTGGCATTACCAGAGACACAAACTCCGTCTTC
GCCTTTTCTTTTGCCTCTAAAATATCTTCCGCCATTATAAAACAGCATGCTCTCACTCCAACTTTTAT
TTATCTACAAACATTAAATC**CACCTG**AACTAGAACAGAAAGAGAGAGAAACTATTATTTCAGCAAACC
ATACCAACAAAAAAGACAGAGATCTTTTAGTTACCTTATCCAGTTTCTTGAAACAGAGTACTCTTCTG
ATCA

>AtCBF4, AB015478, 1380 bp, 2 MYC.

AAACAAACCTTTTGTCTTGTATCCGTTGGTGCCATCGGACGAGTTGACCCGCCACCGCCACCGAGTCT
GTTGAAGACTCGTTTGTTTCCTCCGTTCAAATCGAAATCCATAGCTTTTTTTATTGCCTCTCACTCTC
TTTCTCTAGATTACAATAACAACAGACGCAAATTAACACAAGACCGAACTTAAAAGTAAGGATTTTTC
CGATTGGTATATAGATAAATCAAGCTGCAAAATTGGAGGAAACCCTAATTACCAAAAGAATCTCTGAG
AAGTAGTAATCAAAATATGAAAGTATGAGAAAGAAAAATTTGGGATTGTTGTGAATGTGGTGAAGTGA
TAATAATGCAGAGAAAAAAAGAGAAAAACTTTGGAGTGGATTTATTTGTATGTGTTTATAAGCTTTGGA
TCTTCGTGGAGAGGACAGAAGTAGTAGAAGAGAAAGAGAAAGTGAACACCAGTCGCGTCAGCTTGACG
TCTATTGTTTAGTTACTATTTTATCCCTATGATGGCTCTATTTTACAAAATTACTTTAGTGTTTTCTC
TCGGTTTAAATCGGTTCTGATTTAGACCGGATTGTAGGTTTATGCTTTTGCCTTGGAATGAAGGGTTC
TAGGCACAAGCACAACCCTATACATAATGAAGGGTTTTTAAATCTAGCGAGCCAAGAAACCCCACACT
TCTATAACTGAAAGAAAAGGACATGTAAAGTTAGTGAGGGCCAGGGACATGTTTTTCTAGTGAAAAGA
TTTC**CAAGTG**TTCAGTGTTTGCTTAGAAGAAAAAAGATAGATGATTACCATAAGGCGTGACTCATGAG
AAGCTTGAGGCCACATATAAATATTCTACCTTTTTTAAGAACCACAATATAGAAGCTATATCAGTCTA
ATGGTCAAGGACCTCTATTACTCATT**CACGTG**TTTGGACTAATGTGTAAGAGATTTTATAGTTTTAAT
TTTTTTTGCTGAGTCTACAAGGAATATATGATAATGTGAATCAGGATTATATTTGGAAAAGCAGAGGT
TTCAAGCTTTTATAAGCTCTGGAAATTTGGTAAAGATTGTCAATACATCAACGAAACAAAGAGAAATC
TATTTACAGTTTCAGATACAACAAAACAAAATTTTCAAATTCTATAGCTCTAATTTGATCAAATCTGT
TTACCAAAAACAAACAAAGAACATTCTAACAAGAATAAATGATAGACGATTTATTACTAAAGATTTGG
TACCAAACAAAAAATACAGGTACGGAACTCCGGACCAATCCTAGATGTCCAAATAATCCCTAAACTAA
CTTATTATCCAGAAAAAGAGCCAAAAAATTACTCGTCAAAACTCCAGAGTGACACGTCACCCACTCCG
TCAAAGTCGTTATGATTCCA

>AtDDF1, AC025417, 1500 bp, 10 MYC.

TTATTATTTAAAAAAAAAAACGAACCTCG**CAGATG**CCATTAACGAGGTAACCAAAGAAGAAACCAGAA
GAA**CAGATG**AGAAATTGTTGC**CATCTG**GGTTTATCAGAAAGAGGAATTCCAAACAGTTTTAT**CATTTG**
TTCTTCGCTGTTAGTCTTCATGATCTTGAAAAAAAGAATAAATCGGACCAAGGAATGTTATGTTAGAA

ATGAAAAAAAAAAACTGAAATCTGATTTTCTTTTTGTTCCTCTCTTTGAAGATTTGGGATTTGTAAAA
GTGAAAGAGTCAGTGATATTGACAAGTAGATGAAGAAAGGTATTAAAT**CAAATG**AAGAAAGAACAAAA
GAATTCTCATTTTGGATTTCTCTTTTTTTTTTTTTCTCTTACTCTCTAACTTAATGCTCACGAACACT
TTTTAGCATTAGTATTGAAGTATTAAGGGTCTAATTGTCTTTTAGTTTTCTAACGCTTATATTTGTAT
TTATCTTAGTATTTATAATTTATTAGATAATTATGTTTATTTAGACCAAAATGTAAAACCTTTTTTTT
TTTTTGTCAAAACCAAAATGTAAAACCTAAAAACCGGTTT**CATTTG**ATGAATTTTATTTCAGAGAATA
TTTTATTTTGAAAAATAAATCCGTCAAGTTAAGTTGGATTTTAGTGAACAATTATATGATTAACACGA
ACTTACTTTGCTTCAAACCTCGTATAACGAACTCTCCCAGTATCGGGACTTGTGTGACTCGAGTTAGA
GCGAATGATGAGCTGGCAAAACAAAAAGTCAGTTTCAGGCAGCCACGTTTCATTTAGTTCCGCGTCGA
TGTTCATGCCACGT**CAGATG**ATGCTTTAGTATTTTTCCATAAC**CAAGTG**TAATAACATGAGAAAATAA
AATAACTAACCTTTTGACGGAGTATTAAAATATTTAATTTGAGGACCTACCATAAATTCCGTGAAAAG
ACCTTACCATGAATTTCGTTTCAAGGTATTCAATTTATTAGTTGTAAACTTGAATCGTGACATGCTGA
GTATTTCGGGACTTGTAATTCCACCGGAAAACTTATTAAATTTGTTTTTTGTAGTCAAGCCTATTATT
ATTAATATAATCATAAGTCGTTGAAAGCACAAATATAAAATAGTATTCTTAAGAGAGTATAACTATAA
TACTCGCAAGTTAAAACTGTATTTTTTTGGT**CAATTG**CTGTCAAGATTTATGTGCTAGTATCCAGCAA
GATTTGATTAAATCCCCTTTAATTCTATTACAACTTACAAGTAGATTACACTGTGAAAAAATTATTTC
TA**CAACTG**TCACGTTGTAGTTACAAATTGGGCCTAGGCTTTGTATATTAAGATTCATAAAGCCCAATA
ATTTTATAATTTTATAGGTACAAAGATTCCATACCATCAAGGAAACTTCTAACGCCTGGGACACATCT
ACAATCTTAAAACTCACATAAATACCTCTTACATAAAAGCTAAATATTATGTATAGTAGATATGTGAG
TCCC

>AtDDF2, AC010795, 1500 bp, 4 MYC.

AATGTATTTGATTTTATATACGAGGGAAACTTGAAGAGTATCGGATTAATTTCGATTAGTTCATATCT
AGTCCACTTAATTAAGGGCATTTCCCTTCAAATATTTTGTCTTTTAAGTTGGTTACTCGCAACGTGGT
TGCGTGAAGGCAGAAAACTTACATTGTCATATACTGTCTTTACATATATGAAATCTTAATAAGAAAGC
AAAAAAGAGATTAAAGTTTTTTTTTTTGTTGTCGAAGATTAAAGTTTTTATATGCATGGAAAGCTAGC
TAATATTTTAAAAGTCGCTTTATAGGGATAAGCTTTCTG**CAAATG**GATTTTTCCGTGAAAGTTATACT
TAATCAAAAAGATCATATCCGTGAGTCGAATACTTAAATAAGACCACGTTTGTCGTTGGTATAGTGAT
CCGTTTGTGACCATTGGATCATAGTCACTGGCAACAACAAAAGAGGAAGAGAAAGTAAAAAAAACATC
AATTTTTTTTTGTAAACTTTGGAATGAGACTTTCGCATGTTAGATAAACTAATGTGGCGGGATTCTTT
TTCCGTTTATTACAAATAAAATATACTCAGAATATTTTCTAACCAAGAAAATTACATTGTATGATTCT
ACTACATGCATCTCGACCGATTCATGGTGATAT**CATATG**ATTTTACTGCATGATCTATCGATCTTAAA
TTCTTAATTAAACTTTATAGCTACTACTGGAAAAAAAAAAAAGATCTTAACCAATAATGGGATTTTTCC
GGTTATGAAAAAAAATATGATTTTTCTTGACACGATAGTACGAATACGATACGTCTTGTTTTTGTGAG
CATGGTATAGTACATACAACTCTGGTCAACATAATATGGGTTACATACATCCCATTATTGTCTGGAAA
TAAAACAGCTTTATTATTTATAAGTGTTTTCTTTTCGGATTTGCCATTTTATTTATCTGAGAAAAATA
TTGAATTCTCCAACACAGCTCAGGAAAGCTGAAAATAGCAAAGAAACAATCAGAAAGAACCCGATGGA
AGAGACAGGGATAGAGAATCCATATTTTCTCAGTAATTAAGAAACAAAAACAAGAAAAGGACAAAACT
CTCGTCTTCTTGTGAATTATGGGCCTTTCAATTTAGACCTTTTATTTAGCC**CAATTG**AGATTTCAAAG
CCCACTATTGAGTGTTCTTCGAATGGAATCTGAACCAAAGATGTTATTTGAGTTGTTGTTTTGTCTCA
TGAGTTATTGACTTCGAGAGAGAGGGTTTCAAAGTTTTTTTTTTTTGGGATGTATATTAGGTAAGAT
TGTCT**CAAATG**GCAACAACATTTTTGATTATTTCTTAATAAAACTTATACTAATACAATTTGGTAAAT
ACCTAGGAAACTAGATAAACCAAAAAATAAAACACATTTAACATATAGGCTGGACCTATCATATCTAT
GCCCATCATTAAGCCCATTTATTATGCATACATACACATACGATCTTTGGTTTTTTAACACATACGTA
TTTT

>*Tm*CBF5, AY951947, (*Ta*CBFII-5.1), 940 bp, 7 MYC.

AAGCTTCAGCATGAATGGCGAAAATG**CACGTG**AAGAAGCTCCGATCGGTAGGACGTAACAGACGGGCA
CAAATGGACAAAGAGTGCTCGACCAACCTATAAACAAACAATAGGCGAAAAAACCACGTATCTGTTTG
TGTCGGCGCGTCGGAGTGCTCTAACCTCTATGCAACAAAGAGTCGCCGCGTTAAGGCGGGCGGGCGGG
CGGCAGG**CACATG**GCGTCCCGCCTCGGCACCTCGTGTCAGCCCGGCAGCCCCGCCACGTACCAAAGAG
CACGCAAAAGGGCAAGGTTAACCTGACGTGCGCCGCGCCACGCCGGGGATCGTCGTTATCGCCGCCCG
GCGTGGGCGCGCAGACGCGCAGT**CAGTTG**ACAGGCCGAAGCACCGTCCGTCCCACGCAGGCACG**CAGC**

TGCTCTGCGGGACAGAGTACCAGTACTAGGATAATAACGGCGGGCCGTGGACGCGTTCGGGCGGGCGG
CGCCAGGGCACGAGCCGCCGCCCATCCGTCCCTGCCG**CACGTG**CTTTCCTCGAGATCCGGAGCTCTAC
CAGTACACCATAGTCTGACCCACTGACACAGTACGATGCCGGCCGGCCAAGACCAGCAGAAAATCCCG
TCTCTGTCGCCGTCTC**CACGTG**GCCTCTCCCCCTTCCGGTCGCCTTGCTTCCGATGCAAAGTGTGCAA
TTCCGAACTCTTCTAGTTGTAGCCTTGTATACTCCGCGCGAAGCTAGCCCGCCACGCCAACGCAGCCG
GCCTCCCTCCGCCACCGTGTCCCGCGACGCGCCGCCCATTCGGACCCGCCACGCGCCCCGGCCGAATC
CTATATACACACGTCGCTCTCCTCGCTCCCTCCCTCCCGATCATACAAACCTCGATCACAAGCCAACA
CCATTGATTCGCTAGCTACAGTGTCTGCAGATAAGCAAACGATCGATCCGTGCAAG
>*Ta*CBFIIIa-6.1-Manitou, 814 bp, 4 MYC.

AAACTGGAGTACAAAAGACAACACCGGCGATAGGGGCTCCGTGGTTTGCGTCTGCTCATCCACGTCAG
AAATGA**CAATTG**AAGTCGGTCACCACAACCCCCACCC**CACGTG**CGTCACCTAGGAGCAAAAACTCCGC
TTCCACTCTTGGTCTCCAACATCGAATGCGAATTCTGATTTTTGGTCCGCAGCCCCGATGGCGAAGAG
ATGATGAGCTTCGACGTTAAGAATTGGACGTCAAACAGCACTGACCATCTAGACCCCATCCAAAAAGG
AATCGTACCACACTGGCACCAGATACCCTCGGCTAAACCAAAACCCGAATTTGCAAACTGCGGCTCAA
AAAAAAAACCATTAACAACGACGAGGAGGTAAAGTAAAACGGCCGCAGGACAGGGCGCAAGGCCG**CAA**
GTGAGCAGGAGCTAGGCGGAAAACGATAACAGGGGCGAAGCGAAAGCAGGAAGGTAGTACTGGTCGAA
GCTGTATTATAATAACAAAGAAAAGAAATCTTGGGGTCACTCTCGAAGTCTCTCACTTTGCCCCC**CAC**
TTGCGTGCGCCTTCGTACGGTGCATCCCCCCTTCGCCGGCCGTATCCAGCCTGTGTCCTCTGTCCTCT
CCTCTCAGCCGCTGGCTCACTCGCCGCGCTTTGCCCCCTGCCTCCCGCCGCGAGCCACCGCCACCTAT
AAGTACGCGCTCCCGAAACCACCCCGCTCAAATTTCTCATCGCAGACCAAAACCACTTCACACTCGAT
CAACCCAAGAGCACCCAGGACCTCCAAGCCGGCCAAGGAACAGAGCATCGAAGGAAGAAGAAGATG
>TmCBF18, AY951946, (*Ta*CBFIIIb-18), 521 bp, 6 MYC.

AAGCTTCTCAGTTTCAGCAAAGCCCTTGC**CAGTTG**CCACCACTATGCCAGTCGCTAGCCTCACGAGTC
ACGCATGG**CACCTG**ACCTGACCTGACTTTCTTCACACTCACTTAGCGAACAGGCAAACGCCCGGCATG
AAAGCGCCATCCTTCC**CACGTG**GTAACATAACGCGCTGCAGGAAAGCGCCACCATGCCATAAATCGAC
GTTTCAGATAGCGACTTAACAA**CACGTG**AACTCAGCTCGCGGCTATGAGCAAGGAAGCTATACCAAGC
CGCGTTTTATCCGCTGGAGTGCACACAAAGCGGCGTAAGCTGCGTTTCCAAAGCGCGTCTTGG**CATAT**
GGCGAGTGGACGCGATGCAGCAACGCGGCTAAGGAAA**CAGCTG**TCATCTTATTGTCACACGTATAAAT
ACGCACACGCACTAGCACTACACCAAGCAACACTCTCACTCTCTCGCTCGAGCACAAGCCTCAAGCTA
GATACCACCAGCGAGCCATCCACCGGCAGCCAGCAGCCAGCAGCC
>*Tm*CBF3, AY951949, (*Ta*CBFIIIc-3.2), 517 bp, 2 MYC.

AAGCTTCAAGAATTAGTTATTTTTACATATAGACCGTGTATTGAAGATGTTCTAAGTGGTGCTCCTTT
GTGCCTTCCGTCCCCCTAGTACTCGGGGAACCAGACGACCCCATGCAGCAGTGACTGCTGCCTTCTCT
TTGCAGCCGAACAGCCGGCGGACCAATCAGTCAAGGCAATCACCGCTGCATTAAGCCAGCACGAAGCT
GCCTTTTTTTGCTTAACACTGCGAAGCCAAAAGCCCCCACACGCCCACCAGGAGAGAAGTCACACGAC
GCTATCACCCCACAGTCCCTTTGTCAC**CAGCTG**TCCGGACACCGCATCCCTCCCGCCGTCCCAAGCGC
GTTCATA**CACTTG**AACCTCCAGCATCACGCATACCTATATATAAGGAAGTATCCCACACTCTCGCTCA
AGCTCAACAAGCTAGCTCACACTCCTCAGTCCTCCGTAAGCTCAAGCAGCAAGCTCGACTGCTCAAGC
AGGAATCCACCAGCCAATCACCCAGCACTCAGCCGGCAGCC
>*Ta*CBFIIIc-3.2- Norstar, 523 bp, 1MYC

ATCCCTAAATTTCAGTTTTTTAGGTATTTTTTACGATTAATTTTTGTATAGGGGCGTGGAGCACCCAG
GAGCTCCTGCGTATTTTCTCATAAACTGTACACCATAGTGGGCAGCCTTTCTGCCTTCGGCCTCCGTA
GACCTGGGGAACCAGACGACCCCATTCAGCAGTGATTGCTGTCTTCTCTTTGCGGCCGATCAGCCCGC
GGACCAATCAGGCAAGGCAATCACCGCTGCATCAACACTGTTAAGCCAGCAGGAAGTTGCCTGTTTTT
GTTTAACACACTGCAAAGCCAAAAGCCCCCACACGCCCACCAgAAgAAAAGTCACACGACGCTATCAC
CCCACGGTCCCTTTGTCAC**CAGCTG**TCCGGACACCGCATCCCTACCGCCGTCCCAAGCGCGTTGATAC
ACTTCAACCTCCAGCATCACGCATACCTATATATAAGTCTCCCACACTCTCGCTCAAACTCAACAAGC
ATCTCACACTCCTCAGCCTCAGTAAGCTCAAGCACCAAGCTCGACTG
>*Tm*CBF10, AY951950, (*Ta*CBFIIIc-B10), 1500 bp, 12 MYC.

GCATCGTGCCTCCACACCGCTCCAACGGAGACGTACTTCCCCTCAAAAGGAAGGAACTTCGGTAACAC
ATCCTCGTCTTCACCGGATCCACTCTTGGTTATCTCTTACCTTTACTTATGCAAGCTCTTTAGTGATT
ATTCCCTTGCTTGCTTGTGTGCTTGTTGTTGTTGCATCATATAGGTTGCTCACCTAGTTGCACATCTA
GACAACCTACTTTGATGCAAAGATTAATTTGGTAAAGAAAAGCTAAAAATTGGTAGTTGCCTATTCAC
CCCCCCTCTAGTCAACCATATCGATCCTTTCAACAACGAGCCGCACGGACGCGGATTTGGTGAACATG
GGTGCACGCGCACCCTTTATAAATAGTAAATTCAAAAAAAATGCTAGAAAAAATAAAAAAAACGAACTT
ATTTTTGTAATATACATACTCAACCGGTATACTCG**CATATG**AAGTTTCATGAAGAAACACATCCGTGG
TAATCTGGGCAAAAATGACAAAATCGAAAGCTAATCGAAGCTATATTAAAAAAGGACTATTTAATGAA
CAGTATGGTCG**CATTTG**TATTTTCTTCACTAAAAATACCATGGATGTCAATACATCATGAAACTTCA**C**
ACGTGAGTAGAATGATCGACTAAGTTTCATACCGCGAAATTTCAGTTTTTTTTTATTTTTTCTAGTAT
TTTTTATGAATTTA**C**TATT**CACGTG**GGTGCGCGCGCATCCATGTTCACCTTTGTATTTTCGCGAGTCG
CACGCCTTGCCTAATTCTTTCTTCTCATTGCTTTAGTGGTTGTGCACGTCCTATCCATGCAAAGGT**CA**
GGTGTGCGTTTAGCTCCTTGAGACACTACAATTCAATGAAAACGCCCTCTGCTTCCTTGAGTCC**CAAT**
TGCTTCGGAC**CAGATG**TCGGTAATTTGGCGCCATCGTGGTTCACGGTCACCGGTGCCTCTCGTGCCTT
CCGCGTCCAGATAACTTGGAGAAACACAAGACCTATTTAATAGTACTCC**CATCTG**TTTCAAAATATAT
TAGTGTCAAGAACGTTCTTATATTGGGGGACGGAGCGAGTAA**CAACTG**TCTGTCCCTTCTAGCCTTCA
CTTCGTCACGAGCCTACCAGCCGTCCAACCCCAGCCTTCACTAGTTCACTGTGTTAAGCAAGCAGGAA
GTTGCCTCTTGTTTAACACTGCAAAGCCGAAAGCCCCCACACGCCCAGCAGGAGAAAAGT**CACATG**AG
GCTATTAACGCCACTTAATTTAATCCTACACGGTCCCATCGTCAC**CAGCTG**TTTGTACGCCGGCATCC
CGACCGCTGTCTTCAGCGCGTTCATACATTCGAAGCTCCAGCACAGCACATAACTATAAATACATCTC
ACACACTCCA**CAGATG**CCAACAGCGAACACTCCCACTCAAGCTCAAGTATCTCGCACCGGAGTAGCTC
ACACTCCTCACTAAGATCAAGCACCAAGCTCGACTGCTCAGAAAGGAAGCGCCGAGCACTCTGCCGGT
TGCC

>*Ta*CBFIIIc-B10-Norstar, 497 bp, 1MYC

ATACGTGGACTGCTGGAATACTCATGGCACCATGGCTGTTCATCCACTAAGTCGTGCCCCTTTCTGCC
TTGAGCCTCCGAAGACCTCGGGGAACCAGACGACCCCACTCAGCAGTGCATGCTGTCTTCTCTTTGCG
GCCGAACAGCCGGCGGACCAATCAGGCAAGGCAATCACCGCTGCATCAACACTGTTAAGCCATCAGGA
AGTTGCCTTTTTTTTTTGTTTAACACTGCAAAGCCAGAAGCCCCCACACGCCCACGGTCCACCAGGAG
AAAAGTCACACGACGCTATCACCCCGCAGTTCCTTTGTCAC**CAGCTG**TCCGGACACTGCATCCCTACC
GCCGTCTCAAGCGCGTTGATACACTTCAACCTCCAGCATCATGCATACCTATATATAAGTATCCCACA
CTCTCGCTCAAGCTCAACAAGCATCTCACAGCTAGCTCACACTCCTCACTGAGCTAAGCACCAACCTC
GACTGCCCAAGGAAGCCACCT

>*Tm*CBF13, AY951951, (*Ta*CBFIIIc-D3), 1266 bp, 7 MYC.

AAGCTTTATATGACAGAACAATAAAGAAGTGGTCAGGCCAGTGGGGGCAAATATCCCGGGCCGAGCGG
AGGCAAAGCCATGTATCGGAAGTCAGGTATAAGTGGGCAAAAAAAT**CAAATG**AGGGCGGACCTGCACG
TAGATTCGCCTCTGCCCCCTATGTGTTTGATACTCTTACTTAGCAAAATATA**CAATTG**ATGTCCTATA
TTTGTGGGTATCAAAATACCTTCACTCTCATGGATAGACTTTAATTTCATCTTCAGCGCCGTCCTTAG
C**CAGATG**ATTAGACGTGCAATTC**CACTTG**TTGAGTACTCCTAACTAGTACGTGTGCTCGATAATTCTT
AAATCTTCTCGAAGGGCAGTCCTTGACCAGATTATTGAGGTGTGGAGGATCTCTTTGCCAAACTCACT
CTCAGGCCTTCTCACCTTTTTTTTCGAAACGGAATCTTTTCACCTTTTAATTTTGAAATAAATTTGGA
GTAGTTTCAGAACGTTGTAATTTGTCTTTGTGGTGTACTGTATACTACTATTCCTGCATCTTCTTTTT
TTTGCAGGGTACTATTCCTGCTTCTGTTGACGGATAATTTGCATTAGTGAACAGAGGGCGAATTTAGA
TGTAGTTTTCATTTCTGCTATTTTATATATCCTTTAGCCTTTTTCAAACAACAACAACATGCGTGTTT
TTAGTCCACTGTCTTCTTTCCGCCGTTACATTTTCCAGGCTAGGTGGCGCTTTCTTGCCTCCAGTTTT
TATTCATTTTGACTGCGTCCTGGAACCATGAAAACCGGAAAGAATAAGGCAATATGCGTGGCTGGGTC
TCCACGGGCTGATCCGGTCGGTGGAGCCCAATTTACCGCTGCCCACCAAAGAAGGAAGCTGCCTTCCT
CTAGTCTAGAAAAAAATGCTGCAAGCTGACTTTGCTTAACACGCTGCAATTCGCCCCCTTAAAAAAGC
ACGCTGCAATTCGCCACGCGTCAATCT**CACATG**AAGTCGCAGTACAACAGGCCCTTAATCGACTAGAA
TAAGGCCACGTAGTAC**CAGCTG**CTGTCCAAAGCGCGTTCATACACGCTTGAAGCTCCAACGCGACACA
TACCTATAAATACGTCCCCACACACTCCA**CAGATG**TTAGTAACACCGATCACCCGACGCTCTCAATCA

AGCTCAAGCATCCCACGGCTAAGCTCACACTCCTCGATAAGCTCAAGCACCAAGCTCGTCTGCTCATG
AAGGACGCCACCAGCCCGTCACCAAGCACTCCGGCGTTCGCC
>*Tm*CBF12, AY951944, (*Ta*CBFIIId-12.12), 1500 bp, 8 MYC.

TCTTCTTGTTGACTTCAACACCAACAAAAAAAAATAAGAATATGACGGGCAACCCATAAATTCCTTGA
TTTAAGTAGTGAAT**CACCTG**TCATGAGGAGGAAGAGAACATGGTGACATGCGGAGGGTTAAGAAGAGA
GATGATCAAGCTGACAGGGTTCATGCA**CAGGTG**CTCGACATGCACCCGCTGTGGGTAATCTCGATGCC
ATCCCCTCT**CACTTG**ATTAACCCTTGTCGATGATGGTTTAT**CAAATG**ATGGAATGGGAGGATGAAACG
CCAAGGAAGAATGTGAGTCTAGCTTGGCATAGCTACAACGAAAAAAACTAGGCAGGCCAAGCTAAAGG
AAAGCACAAAACAAAAATCAAGTTATAAACCATATATGACTTCCGTAAGCATTGGTAAGCACTAACTA
CTTTTGGTTCTTATTGGTTGCAAGATCAA**CAGATG**ACGCCAGTGAAATTTGGGAAAAGAATCAAAGAG
CGGAAATTACACCTATGCTTGTTGGCTAGGTCGCCCATTGCCATCTCAAAGGGGATGGATGCGGATTG
TTTATATGTTTCCGATACAATTTGTTGAAATTCCTTCTAGTCCGGATACAAGGTTGTCTTGGTATGCT
AGTAATTATATTTACTGATGGGCTAAGTGGACCATGACACACCTCGCCCCTCCAGTATCCCCGAGTGA
CTGTACCTACACGCCTCTGTCGTTTTCTTTGCTTTAGATGATTTGGCAAGCTTGAGGGACTAATTGCG
TTCGTAATAGAAAAGCAAAAGTCCTTCATAGATCAGGGTTGGTTTTTGTTTGATACCCGCTACTCTAC
CCGAAAAGCACATAAAAGTTACTCCCTTCTTAAACTAATATAAAAACATTTAGAAAACTAAAATAATA
ATCTAAATACTGTTATATTAGTTTACAGAAGGAGTA**CAACTG**ATTCAAACAAGGA**CAAATG**TTGT**CAC**
ATGGTAGCTAAATACACTTCACCAGCAAAACAAAACGTAAATACACTTTCACATCACAAGAACACACT
GATTTGATGCTAAGATGAAAACGCCGGGAATGGGATACACCAACACACCTAAATTAGCAAGTTGGCCG
TCACCCACGTTTTCTCGCACCCAGCCGACGAGCACAAAACGCGGTTAAAAAGATACTCCATGCATGCC
TATGGGCTGTGAGGTGATGGAGGGAAGCTGCCTTTTCTTTGCTTAACGCTGCAAACCCCCTACACGTC
TACACCGAGGAGTCATGCTACACGCGTCGACCGGGGAGCGGTCCAGCGCCGCACACCACAAGCTGTCC
CCGCGCCGACGGGCAAAGCCGTCGCATTTCACCGCGTACCGCAAAACTATAAATAAGCACACTCTAGG
CAAGCCACTGGTTTACTGAGTACTCCTAGTCTCATAGGTCATAGCCACCAAGGCAGCAGCCATCACGG
CTCCTTACTCCAGCCAGCCAACTAGCCTAGCAAAAGTACATACCACCGCCTAGCTTGCTCGGCGGAGC
GGCA
>*Tm*CBF15, AY951944, (*Ta*CBFIIId-15.2) 1500 bp, 8 MYC.

CATTGCAGGCTCCACCGCCTCACCACCCACACCTCCACCCCGCAGCCACACACAGCAGTCGGCCACCC
ACCTCCGTCTCGCAGTCACATACGCCCTCACCAGTTTGACTGTTATGTCCAAGC**CATTTG**TGGATGAA
CGAAGCTGATCTTGGGGTGTGCGCGTGAAGACCACCTCAAGTTCTGCAAGATGGTGGCTGTGTACATA
GTTCACCATAGTTCACCAACGTCAACGAGCTCATCAAGGAGGCGTGACCTGTGATGTCTGTAGTGGCG
AGACGGAGCGGAAGCGCGTGATCCCTGGTGATGATGTTCGAGCGCTTCCAATGAACATCTTCCTTATG
GGCGATGGAGGAGGCCGAGTGATTCCGGACAGAGAGCTCGATGGCACAAGCACCGGTCTCTCGTTCGG
GACCTAGTGGCTGATGCCACTAGTTCAAAATCGAGCTTTGGAGGCAGTTCTCGTACATGGACAAGCAA
CAGTTTCCAGTTCCGTCTCAAAGCTGGTGTGGCTAATACTTCGCATGATGGTTCTGGTTGTGGAGACG
CCTCCAATACCACGGACAACGTTCGTACTAATAAACGCCTCCATTT**T**GAATGCCAGGGGTCTAAATGT
TTCCTTTTTGGTATCCGCCCTGCTAAAACAGT**CAGATG**AAACACTCCTTGATTTCCGTATTTCCTAAT
TGGTCACGCGAATGAAAATGAAAGTAAACATTAAAATAGAGATGTACATACAACATT**CATTTG**CATAG
AGCTCAATCTTTTAGAAGTGAACGCGCATCATCAGAATTCCTGATAAAAATCAAGGATAGTGTGAGTG
TGAGGACTTGAACACTGGTGAGTTGGGGATAACACTGTGCTCGGCCATCTACGGTTGTAATGTGCTTG
AAAAAAAATGTGGTACTTGTCTTCGCACAGAAGTACATAACACGAGGAGAATGCGTTACTTGTCCAAG
GCTGGTCCATGCTAGTGGGCCCCATTTATCAGTTTCAAGCAACACCTTTTCACGGTGCCAATCCATGA
CGCGT**CACTTG**TCCCCTGCCGCTTGCCCACACCGCGTATAACCAGTCTCCAGCAAGAAGATGCCTTCC
TTTTTAAAGCTGCAATTCCCTGACGCGT**CACTTG**TCCCGTGCCGCTTCCCCACACCGCGTACACTCCG
CAAAGACACTCAATCGCTGGAGAACAACATATAAACGCGGCAGAAAAATGCGT**CACTTG**TGCCAGCCT
GCTGTTTGCTAGATGGTCACACGCAAGGGCCC**CAACTG**TCAGTCTCATGAAGCAAGAAGGTGCCTTCA
TTTCAACGCTGCAAATCCCTGACGCGTCTTTCGTCGCTGCCGCTAGTCTGGAGCGCGTCAGGGAGAAC
TATAAAGCAGCGCACTCCACAAGACACTCAAGCATAACCAACACTCCTCAGCTCAAGCTCAGCCTGCT
CTCACACTCCATAGATCGACCTCGGCCCACCGTACGTAAGCTCGAGCCAGCACCGAC**CACCTG**CAGTA
GCA

>TmCBF16, AY951944, (*Ta*CBFIIId-16) 1500 bp, 6 MYC.

CACACCCAGTTCCGTCTCAACCATAGTGGAAAAACGGAAAATATGGTGAATACATGAGAATTCAAAGC
TTGGTTCAACGCTACCTTTCAAAAATGCACTAATCATCAGGTCTGAACTTGACCTATGTCTAAATGAA
GTAAATTATCTTGATTGAGCAATGTATCACATTATTACAATATTAAAAAGGTCTTTTT**CAAGTG**AGAT
GCGCCCCATTTTACGATTTCGCACAGGACCCCGTATTTTGTCGGCCCGGCCCTGTTGTGAGTCATCTA
CTTACGTTATTATCTTAAAACATATACGGTTTTGCTAAAGCACAACTAGATGTGCCCTAAGTATTGCG
CAGCTAAGTCTTGTGACATTGATTTTACGCGAAGATTCGCATGAGTATTTTCTTCTATTTATGTTTGA
TTGAGTCACTTAGATGTGTTTTTAGGACACCTCTAGATATGTCCTAGACATACTCTAGAACATATACT
CCAGTTCGGCCGTCACAAAAGGGCCACTCCAAACGTGTGTGGAGGCTACTGATGAGGCAATTAAAACA
TTGATGGCTCCTTTGATTAAAAGAAATTCTATAGGATTTTTGGAGGATTAAAATTCTTAAGAGTTCTT
TCTATGTTGGT**CATTTG**ATT**CATATG**ACTGAATCTCATAGGAAATTTTCCTATGAAATCTTTTGTGTT
ACATTTCATAGGAAAATTAACATTCACTCGAACCTCTTTTACAATTTGTTGGTTTTTCCTGTGGCATC
AAAACACTCATTGCTAATTTCATAGGATC**CAAGTG**GGCATGTAACTCCAATCCTATACTTTTTTTATT
CCCACGTTTTCGAAATCCCGCCAATCAAAGAGGCCCAGAGTGTTCTTTTGCTTGTGTACGCTCAATCT
GGCCGCTTGTGATTGCAAAAGTAGGTGTCATATTATGTACAGTAACATTCTTGTTTCTCTGTGTTAAA
CATCTTTTTTGCAGGGTCTGTGTTTTAAATATATTGTACACGAGACTGTAGTAATACGATTAGAGGGT
TAGTAAAGGCGACTAACATTAACAAGAATAATGCTACACATACATAGGTTTATGGGGGTTTTACAGGC
AGGATGTTTTTGATTGAAGATTAAGGGGGAGGAGGGGTCTACCCTCGTGAAAATCAAAGGGATGGTTA
GCTAGAAAGAAAGAATTATAGCCACTGTAATTACGTACAACTCTTTGTATATCTAGCATTATTGAATT
AACAATTATCCATCTTGGTCAGGGAGCTAAACCTGGGGTTTAAACCGGTAATAACCAACTT**CACCTG**C
GCTCCTTGTCCAATGTCCTTGGAGCTTTCTTGCCTCTGCTAGAGGACAAGATAACGCGGGAAAAAGGT
CTGTCTCCCTGGGCCTCCATGGGCCAAGCTAGTGGAGCCCAATCATCGCCCTTGAGCACGAGAAGGAT
CCACCACTAAAAAAGGAGCACGAGAAGGATCCTTTGGATTATCTCCGCAGACCCCTCACGCGT**CACTT**
GTCG

>*Tm*CBF17, AY951945, (*Tm*CBFIIId-17), 1500 bp, 3 MYC.

GGTGGTTTTGGACAATGACGATGAGGAGGAGGAAGAGGAGGAGGCAGCGGCGAATATTGAATCTGCTG
ACTCGTTGGATGAGCCAGATCTGCCGCAAGTACTCAGGGATGATGATCAAGTATATCTATTGACCGAT
GAAGATATACAGCTCGTGAGTGCTGCATTTCCTAATGATGCACGGCGGTTTCAGGAAGCCCAGGTACG
TATCGGTTACTTCCCTTCTTCCCTGTGGATGAGAAGTAAACT**CATATG**ATGATGCTTCGTGGACATTA
TACTGAGCTA**CATTTG**CATTATATCGCATTTATCTCTTGAACACTGGCTAAAATATCCACTGAATTGG
TTGCAGAGGCTGAAAGATGCAAAATCAAGATCACGGAAATCCAAGCTGAGCCTAGCAGATAGCATGTT
GGAGACACCGAAGGGTCCAAGACCCAGCGGGGTTCAGCTCAGCATCAAGGAGTTTTACCGGTCGAAGA
AAGCGGCGGGAGACGAAGCCGGAAAGAAGCCGGTCGTCGAGGGTGAGTCGTCGTCGTCGAGAGCTGGC
TCGAGAAAATCACCGCCGGTGGACCTGACAAAAAGGATGCCGAAGTCCCTGCGACGACGCTTGCTCTT
CGACTGAGCAGCCCGTGGTTGGGTTCTCCTGACGGTACTCTTGCTGCCACTGCTGTACATGCTCTCGG
TGAATGTTTTAGTGTTTCTCTTGGCCCCTTAGATGTAAATTAGTTATGTAGTCCATCTTCTGTATAGT
GAAGACACGATCTTATTATTGTGGAGGAAAGAGAGCAGGCTTTTGTTTAATTTAACTGAGTGTTGCAG
CCTCTTGTGGTGTGTGCTGTGATTACCAGGC**CAGATG**CTGGTGTTGATGCTATTTTTGTAGAATTACA
GCATGTAGAGCCCTTTTTTTAGTGGACTTCGCTTGATTTGAACGACTAGAATTCATGTAAATTTGCAT
AGTCTTTTTAATGTTTGACCTGAAGAAATTATGTAAATTTGCATGGTCTCTTTAATGTTTGACTTAAA
ATTTTCGGAATATTATATGCTGTACTTCTTTTTTCTAAGGGTAAAATTACATACTTGTGGTCACTCTT
GGCATGGCCTCTGTGCAGTGGATAATTAGGACGACTGAAAATCCGGAAGGCTGTGCGGTTCGAATGAA
CCCAAAAGAACGCGTGTTGAAAGGAAGAAAAGGGCTTTCCCTCTATTTGGGCCTCCATAGGCTCGGCT
AGTGGAGCCCAATCCACTCGTCGAGTGAATGTTGAGAACGAAGCTGCCGGACTGGCTTAACGCTGCGA
GACGCCACGCGTCAGCTTGTCGCCGGCCCGCAAAGGCCGTCGGCGGCGGCCCCACGCATGCTCCGAGC
GCGTCACGCAAAACTATATAAAAGCGCACTCCACACAAGACACTGAGTTATCCATCCACGGACCAGTA
TAATCCCAAGCTAAAGCGAGCACCGAGCTAGCTAGTCACTCACCGTTCACCGATCGACCGGCAGCAGC
GCAG

>*Ta*CBFIIId-A19-Norstar, 1446 bp, 9 MYC.

AAAAGGGACCGCCTGACGTTAGATACCTGAAAAAAAATTGGTGTAAGTAAAAAATCTAAGTTGTTGGA
TGCTAATAATGGCCGGATGTTCTTCTTCAATCTCATCCATTTTTCTTCTCGCCCACGAACAGGTCACG
CACGATCCTCCTACCTCCGACCTACCCCACCCCAAC**CACCTG**CTTCCATCCCCCGCGGCCGGAAGGCA
CTGCGGCGCGCTGTCTCGCCCTTCCCTCAACCTTACCCCATAGATGCTGCTTGTCACTTACCATCTTC
TACCTCTCTCCCCCCTCCCCCCACCCAC**CATCTG**CAGCAACTCCGGTGACTCCCACACCCCTAATCTG
CACCCATCAAGTTTCGGGCTCACTCCCGTTGGCTGCCACCGACTCC**CACTTG**TCCACGGCTCGACCTC
GCCATGGTCGACTCTTGCCTCAGTCGTTCCTCACGTCAGTCCCGCCATCGACCTCGTGCGAAGCAGAA
AAGTGACTCACACTAGAAAAATTTCTAGAGACCAAAGATTTAGCAGCCACGAAAAGGTAAAGGCCATT
TTACTCCGTCGGCCAATTCGATATACAACACAAAACAAGTCTACAACA**CACTTG**TCGCCGCACAAAGC
AATCGCAACAAAAGGTACGCGTATCATTCTATCACGTATCACGCAGCGGAAACGGGATGGGTAACTTT
CGTCTGTTCGATCGTGCGTGCGGGGCCAACCGATCGAGCAGGTCGAATCCCTTGTTTCCTTGTCAAAA
CATGGT**CACATG**CCAATCTTGTAAAAGCTAGCGGGACGTTGAAACTTCTTCATATTTCCTGTCAAAAA
TCCTTGCAGCGATGCGATATGTCGCGTC**CATCTG**GGTGATTCGGAAGCAGCCGCTCTGTGGCTTGTAC
GGCCGTACAGTGTCTCCATAGATCGTGTGGCCAAGAGGGGAGAAGCGCATATAAACGCGACGTGTCGC
CTTAAATGGTTAGATATTTGTAGCGCATAGTGCCATCGGTCGACTAGAAGCACTGTT**CATCTG**CCAGT
CGTGGTGGGGATGAACGTAGGTATAGACGGAGCTTAGCTGCGTCGACATGTCATCATCGTACAGAGAT
CGCCATCACCGTGGAACCCAAACGCCCCACCGCCAATCGTCGACCAGGAACACCAGCCAAAATATCAG
GCGCATTCTGCTTCCATGAC**CAGCTG**CCGCGTGCCCCCTCGCCGACGGACGTGCGCCACCC**CATCTG**CC
CTGCCGCGTCACGCATAACTATAAATGAGCGCACTCTGCACACCGCACCCAACTACCCAACACCTCAA
CCACACTCTGCACACCACACTCGTCTATCCAACA
CGTACCTCAAGCTCAAGCTCAAGCCAGCCATACCGCGAAAAAAAAAAGAGAGGAAAAAGCTCAAGCTCA
GCTAGAGTGACCACATCCCCCCAAGCGATCGGCACTCCGACAGCTAGCAATG
>*Tm*CBF2, AY951945, (*Ta*CBFIVa-A), 1500 bp, 5 MYC.

AATGGCCTGCAAGGGTGCGTACGCACGGT**CAGATG**CGGCAGCCGCCAACAAGAAACAAGACCTGAGGC
GATATCCGACGAATCGATCGAGTTGGTAGCAGCAACGATTCACCTCGATTCCGGCAACGACATCAATC
GGCGAAAGGCTCAGACTCAGGCAAAAGGGCAGCGAGTGCCTAGGGAGGTGACGAGGTAGCGGCGACGG
CGGTGGATAGTCATCGGGCGGTGCGTCTGCGACGAAGCAATGGGTGGGGAGGTGGTAAAATGCCCGCA
AAGTTTACTCTATGCGGGAACACATCGACCATGTCCAGTAGGTCGCCTCCTCCACTGCCCTTCTTAAC
GGATTTGCATGCAATCTCTCGGAGTAGCGGCC**CAGGTG**GAAATCAACAAAAACATGCTGTTATATACA
TGTTTGTTGCTGATTTTTGATACTCGCATGGATGTTTGTTGATGATTGTTGTTGTATACATGTTTGTG
ATCTATTAAGAAAAGAAAACAGTGC**CAATTG**CTTCATTTTGCAATGTAATTTTCTATAATATTCAAGT
ACTATGTAACAAATTCCTCCAAAAAATTGCACCTATGGAATTTAAAATTCCACCGATTAATGGTCAAC
CAATTCGTCCAGTTAATGGGCCAATTCACCGATTAATCGCTACTCTCAAGCCGAGCGTGCAGCTAC**CA**
GTTGCCGATTTCCTCAACATTGGGGAGTGCCAGTGAGGTGAGATCCACCCCGGGGATGGCCAGATTCA
GTGACCTGCCCGCCAGATCGTTGGCGACAACATCGGAGCAAGCGCGGGGCAGCGACTGGAAACAATGG
TTGGAGAGGAGAGGGGGAGGATGGAGAGGAAGATGTCAGCAGTGGAGAAGTTCCGCTGGCGCCGTCGG
ACGTTCGCTGGCGATGGCGCCGAGGAGTGAGGGGCGGAGGTAGACTCGACTGCTAGGGTTTATGGTCC
CTGTTGTCGCTATTTTTGTTTGCTAGGCTTAACTATTCAATTTGTCATATTGG**CAAGTG**AATTGCAGT
TTTTAATACCTGGATGGTTGAATGTATCAAGCTGGAAATGTGGTAGCTCACACTCCAACACAAAATCT
TTCTTCATTAGAAAAAGCTGCAGGATAGAACTTTAGCACAAGCATGCTTGTCTTTGACTTGTGAACTG
CTGGACGATTTTGTTGTCACCGCAATGAGATTTCATGTTAATGCCATAAATAGGATGGGACGATCCAG
GTCTTTGATTTAAGTGAGATTAGACGGCAGACCAATCGAGGGGCCCTCGATCGGCCGGCGTGTAGCTC
TGCAATACTAAAAGAGTACGTGAAACTACCTGCTGATGCGTACACAGCTCCACTGCACCTTGCCATTG
GTTCATTGACATGCGGGTCACTGGTCTCCACTCCAGTTATATATACGTCCCATCCTCCTCTCAAGCTC
CACCACCTCTCCGGCTCACTAGCTCCGAACACACAACTACCTCGCTATACCACAACCCGCTCTCGACG
AGTG
>*Ta*CBFIVb-D20-Norstar, 1090 bp, 7 MYC.

AAAATTTGAAC**CATGTG**TTCACAATTAAATTTCCATTTTTCTTTTTTGGAATTTTTATGGACAATGTC
CGCGTTGTAAGGAGAAATTTTGAGCTTATTGATGAAGGATCAATTCATG**CAACTG**TGTCTCATGTAGC

TTAAAGATGCCTTCGTAAAACTGCGGTATTTACTTTCTCTGTAAAGAAATATAATAGCGTTTAGATTA
CCAGAGTAGTGATCTGAAGGCTCTTATATTTTTTTATGGAGGGAGTACTTTTTTTCAATCTCATGTTC
ACACTTTTTTTTGAAATTATAGTAGATTTTTATTCAAAGCAAAAAAAGAAAGTAAAACGGCAGGCCAA
CGGACAGCAGTAACGACAACCGCTGCAGTACGTATACTTACGAAATGGTTCTAAAAATAGGCGCTATT
CGTTCACGAGCGGTGACTATATCGGGCAGCTTAATCCACTCCACTGGCTGGTGTAGACGTACGCTACG
CTAATTAATCCACAAGATGTGGTAGGGAATCCATCTACTAGTCACTAGTCGGCCGGTTGAGGCCACGT
TGGGGCCCTGGATCAGAGATCTTTTCACTGCTTCTGTGCTTCACGTCGTGATCACGAGCTGTCCCTTC
CCCTCCGTCCTCTGAATCCTATGGC**CAACTG**GCACGATGCC**CAGATG**TGATGCGTTTGTCCT**CACTTG**
GTTAGATATGTGGCTCCATCCCCTAAATAAAAGAAAGATATGTGGCTCCATTATGCTAGTTATTCTTC
GCACGGCTA**CACTTG**TAGCAACGACGGGTAAAACGTACTACAAGCAAACAAAAGTGATCCCCTGGGGG
CACATAAGCCTTTG**CT**TGCTTGCAAAGTAAAAAGGAGCGAGAGCAAAAGGAAGGAGATAGAGATGGAG
GGTGGGACCCTACCGCTTTGTCCGTCTCCTTCTGTGGCGCGCCGCTGCCACCACCCGAGCGCCCCGCC
GGGTAACCACTAGACCACCTTCCCTATTTATTCTTCCCCGCCTCGGTTCCCTTCTTCTCATCCCAACT
CGACCGCTCAAACCAACCTGCAACTCTCAACGCAGCGACTTTCCACTAGTTTTTGACGCTG**CAACTG**A
TG
>*Tm*CBF14, AY951948, (*Ta*CBFCBFIVc-14.1), 624 bp, 2 MYC.

AAGCTTTGAACTACCGCAGCCCTGGATCTTTTCCCGATTTCATGTCACAAAAATCCTTGCCGATATAA
GCGAGCTGTCCTTGT**CATTTG**CTCACTCAGTCGACAGTGGGTGAGTGTGAGTGTGAGTGTGAGGAAGC
AGGCGCTGCGCTCTCACGCGCCCGGTGACACAGTTACGTCGCCCTCTGGGCCTCCAACAGGGCCCCGA
AAGATATGCCAAGATGCTTATTTTATTCATTAATTTTCGCCACAAAAGATGCTTGGGTGTGAAGTGTC
GCCTCAACTAAGATACTTGTTCGACGGAAGAAAAGAAAAAGATCTAGTGCTTGCTAGTATGCTTGTCC
TTG**CATGTG**GTGCAGATTATGTGGCCAGATATTGCTCCACGTAGGCGTTGAAGCTAATGTTACTGAGC
GCGTGACCCCACCACTCCCCGCGCTAATCCCCTTCCGCGCTACGGCGAGCAACCACCACTCAGCTATA
TATATACTCCGCCGTTTCCTCATCCAACCAAGCCTCACCGCTCCTCCTCTCCAGCATCCATCTCTCTC
AAATCTCTCAACGCAGCAGCTAAACTCGCTGCTTAATTACCCCACAGTCGGCAGGCTCCCGGCGACAC
TGCGATCGATCG
>*Ta*CBFIVd-4.1, 806 bp, 5 MYC.

ATCATCCTTCTTCAACACACACTACACAAGATTATCCCTTTGGAACGAACTAAGTCAGGCAAGCTTTT
ATCCTCCAACTCTCCAAGTTGTTGTCTAGCTTAACCAACTAGATCGAAGGGTATATCTTGTAAGATGA
CAGGCACACCATTATCATCAACTCGTTTTTTATAAGAGTAGAAATTATAAGATGACAAGCCACTACTA
GTGGA**CATGTG**GACACGGTACAGGTTTCCAGTTTCTTCCAGCGAATAGTGGTGG**CAGTTG**GCGTGTCG
TTTTCGTGACCATAAGAGGCAAGCAGGCACAACTTTGATTGCACGATATAATACCTATCTATGTCAAG
ATCATCCATCGCTCACCCAGTGGACGCCGCTGAGTGTGAGGAAG**CAGGTG**GCTCTCTGGGCCTCCTTC
AAAAAATCCTATGGC**CAAGTG**TGGTTGATGCCAGGATATTTCTTCCTCTTCTTGTGGCGACCATGGTG
ACAACACGGCAAGAATGAGTGAGATATTTCAGTGGCCAGTGCTGACTGTCCTTCAGAAAACGACCTTT
GCCTGGGGGAGGACGTTTGCTAAAAATAGAAATGAGAGGTCCCTAAAAGATAAAATAAAAGAAAGGAC
GTTTGCTAAAAATAGAAAAAGAAGTGGCCAGTGCCTCGCGTGTGTCAGCGCGAGTGCATTTCCACCCG
AGTGACGCGCCGGGTAGCCGCCGCGCTCTCGCCTATATATGCTTGCCCCGCCGCACTCCCATTCTCTC
AGTAAAAACACCATCGCTCAACGTGCTCAACCATCTCCACTAGCTCTCGACT**CAGATG**
>TmCBF4, AY951945, (TaCBFIVd-4.1), 1500 bp, 7 MYC.

GCCGCGGAACCTCGGGGATCTTATTCCGGGGTCATGTCTAAATCCTGGTGAGCTATCTGTGGAGGCCG
GTGAGTGTGAGCCGTGAGGAAGCATACACTCATCACGCTCCGCTCCCTCCGGCCTCTCCAACCTAAAT
TAATCCGGTGGC**CAAGTG**GCAAAATGCCAGGATATTTTTCTTCCTTTTTCGGGAAACAATGATGCTGG
GTTGACATGGAGTGATGTGACGGCGACGAGGAAATAGTGACTAAAAATTACTACTCCCTCTGTCCGAA
AATACTTGTCATTGAGATGGGTGTATCTAGATGTATTTTAGTTCTAGATACATTCATTTTTATCCATT
TTGATGACAAGTATTTTCGGACAGAGGGAGTACTATTCTCTTTTGCCCATATTAATTGACGCTGTTTA
ACACAACTTTATACTGCACTAGAGCAGCGTAAATTAATATGGATTGGATGGAGTAACAAAATACAACT
CTGGTAGGCCGGTGAGAGAGTCGCGAGGAAG**CAGGTG**CTGAGTGCACTTATCACCCCGGCCCTTTCCA
TCTTAGATTAATCATAGAAGAATCTTAGATTAAT**CATGTG**GCCAGTGGTGAGATGACAGGCTGG**CAAG**
TGATGTGCCCCTTAAGGTAGTACTCCCTCGAAGTTAAAAACGCTTTTATATTACGGAACGGAGGAAGT

AGTAGATATTTGTATCATGGACGACTGTTTCTAAAAAAAAATGGGTCTTGTCTTTCTGATAAAATCCA
TAGAAGAAAAGGATGAATCGAAACATTGCGACCACGGCAAAACTCGGGACAAAGTACCTATGTGGTTG
TATTACGTATACTTACATGTTGTGCAATGGGTACATCATTCATCCGGGTTTATAAGTTTCTTCATATT
TTGCGTTGAACTTTTGACCAAAAATTTGACTAGGAAGACACAAGTCACAAAAAATATATTATTGAATT
TGTATTTGAAAGGAAGTCTACAATGATGATATTATTTTTAGTATTATAAACCGACCTTTAAAAATATA
GAGCTATCTGAAAAATTGGACGCAAGTTGTTAACTGAGTTACTACCAAAACAAAGAAAATTACGCGTG
AGTGGATATGAACGAAAGAGCTTAGACCTTAGAGGAGAGGGGTA**CACGTG**TAGCTCGGTGAGACAAAG
GACGTGGAGCACACGCGCTTTTGCTTGCGGATGCGATG**CAGGTG**GGTGATTAACCCTGGCTGGCCGGT
TACCCGCCACCGCCAGTAATCATCCAACTAAACAAGCGGTGAGTGTTAGCAGAAATTACTATAATAAA
CCAAGAGACGAAGTGGAACGCGCGCCATGAAGCTGCCTCCTTCTGCTGGCGCCAGGTAACCACCTACG
ACTTGCTTATATATACTCCCCGTCTCAGTTCTCATTTTTCTCAGACAAAACCTCCACCGATCCATCGA
TCAAACCTCTCTACA**CAGCTG**CTGATTCTTCCAGTACTCCACACTTACAAGCAGCAGAGATCCTCAAC
CTAG

> *Ta*CBFIVb-B9-Norstar, 1546 bp, 14 MYC.

CCTCCAGCGCCACAGCATTCGGCGTCTGCTGTTGCGCTGCCTGGCGCCGCACGCCGCCATCTCTCCCT
GCTCGCGTCTCTCGGCCTCCTCCATGTTCTCCCACTA**CATCTG**GCGCTCCGCCTTCGCCCGCTCCATT
GCCAACGCACGTC**CACATG**TAGCTCTGGAGGGACAACGGCCGGGGCGGTAGCGGTGGTGGTGGCGGTG
GTGGTGCCTGGCCACCGGATGCGGACTCCTCGAGTGCCTGCTCCAGGCCAGGC**CAAATG**GCATTCTTC
ATGCACACGGATTCCCCCATGACACGTCGCGTCGCCTCCTCGAGGGCCGCTTGGAATGCGGCCCAATC
CTCCTCCCGCACCACGAACGCCGATGGTGCAGGGCGGGCATC**CATCTG**GACGACCCCACGATGGCGCA
CGTAGTACTCTGTCGCGACCTAGAAGTCCCAGTTCGGGGACTCCGGCGCGTATGTGGGGTCGCGTCGT
TGCTCCGACGAGAGAAGGGCATGGTGGTTGCGAATCTCCAGTGCTCGCGCGCTCGGACAACGGTATGG
CCGGCATCAGAATG**CATCTG**GGGTCGAGGTGCTAGCCGTGCGGGAGTTTTATGTCCGGCACTCCCAGT
GGTAGTGCACACGGTGCACCGAGATGTGTAGC**CAGGTG**CGTTGGCATCCCGGCCAGTCCACCGTGGCA
CCAAAGTGAAGCTGGCTGGGCGGCGCGATGAGCCGGCCTCGTGGTCGTGCTTGCGACGTTTTTTTTGC
CA**CAGCTG**AAGAAGCCCATGGCGATGACAGCAAGCTAGGGTTCGTGTTGGGCGGCGGAGGAGCAAACT
CTA**CATGTG**GACTTGTTTGGCGAGAAGTGGATGAGCCACCCTCCCCTCCTATTTAAAAAGGACGTGGC
ATGGACACTTCTAGTGGTTGACGTGTGGGCCCAAGATGAATGGTTGGAGCCGACATGCGGCCCGAGTC
AGACTCCGGGAGGACGCGAGACGTGTTTGTCTCGCGTC**CATGTG**GACGCAAATTTTAGCCAGAAATAC
ATTCAAACAGACACCAAACAGACCAAATCTACCCATGCATGATGCGTTGGGCCGTTGCTTCTGTCTGT
TTCGATCCAAACCGATGCACGCGGAG**CAATTG**TGTCGGAGTTGGCCCTATGTGTATGTATGTATTTGT
GAACTAGAGTATGTTGTTTAAACAAAATTAAACAAAAGACGTAAACGCACCACCCATCACATCTCATC
CAGCTAAAACTTGGAAAGCAAC**CAAGTG**ATGCGCGTCGCGGAAATGATAGTAAAA**CAAATG**GGGGCGC
GCCATGAAGCTGCCTTGTGCTCCGGCG**CAAGTG**TCAGCGTGAGTGACACATACACATACTCCTTCTCT
GC**CACTTG**CCATTTCCAGGAAACCACCGCTTGCCTATGCCTATATATGCTCCCCGTCTCCATTCCCCT
TCTTCTCAAACAAGCTCCATCCGCCGATCGATCAAAACCTCTCAACACAGCCGCTGATTCCATCACTA
CTACCACTCCACACCTCTCACGAGCATCTCCGCCAGCTCTCGACTCAATG

>*Ta*CBFIVd-D22-Norstar, 633 bp, 2 MYC.

CTGGCGAGAGCCAATTTGGGATATTTCCTTTTTTGGGGTCTCCTTTGCGTGCCTTGATCAGTGGGAAA
GCTTCCAACTTCTGCTCGGATATTCTTAGGCACCGAAGAAGCGCTTACGAAAATGACAAACCTCTCCA
AAAATAGGTTCTATCCGTTAATGTGCGGTGACCATAAAAGAGGCAAGCAGGCACAACCTTGATTGGAT
CACATCGTATACTACATATTATATGGATGTGGTCCAAAGTACGCCAATTATATGCGGAAAATATGTGG
TCGGCAGTCGACCCCGCGTCCACATAAACAACCAATCTCATGTCGAAATCCTCGTCGCTAAGCGAGGT
ATCGGTCCCGGCTCCGTGGGTCCTTGTGAAATCCTATGGC**CAAGTG**GCAAGATGATAGACAGCCAAGA
TATTTGAGTGGCATTGCTGACTGTACACCGGAAAAGGACGCGCGCCAGGAACCTGCCGCGCTGCTT**CA**
TGTGAGCGTGAGGGTGTGAGTGACGCGCCTATATAGGCTTCCCCCGCCGCGTCCCCTTCTTCTCAGTT
CTCACCCACCCAAACACCGCCGCTGAAACCTCCGGTTCTCAGCACACAGCGACCTGCTCAACCGTCTC
CACTAGCTCTCGACCAAGATG

CONCLUSIONS

Functional genomics is a powerful tool to identify stress-responsive genes and dissect abiotic stress pathways in wheat and other cereals. The goal of my research project is to identify and characterize the important sequences of wheat genes and determine their function. This information will be used by the breeders to improve the important agronomic traits. In the course of my project I focused on the identification and characterization of the AP2 family in wheat.

In the beginning of the project, very few wheat AP2 EST and genes coding for AP2 transcription factor were identified in wheat but to a much lesser extent their functions. Results presented in this study show that more than different AP2 binding factors are encoded by the wheat genomes. However, considering that the size of wheat genomes that are estimated to be 40 times larger than that of the rice genome, which encodes for at least 174 different AP2 binding factors, and 120 times larger than *Arabidopsis* genome, which encodes for 146 different AP2 binding factors, it is expected that the number of AP2 genes present in wheat to be higher than that of the other plant genomes. From genomes sequenced, to date, there is no evidence that a correlation exists between the size and the number of genes in a genome. The challenge following the identification of the AP2 genes is to determine their biological functions.

Based on previous literature and results presented here it is possible to conclude the following conclusions:

AP2 transcription factors found in wheat constitute a major and an important transcription factor family whose sequences and functions are conserved at the angiospermes. Only certain genes would have evolved/moved or specialized after the divergence between species monocotyledones and dicotyledones. This could be explained by the natural selection and by the adaptation of wheat to its environments.

I identified the genes that influence the LT-tolerance capacity of wheat, as a basis for the Triticeae. In this study, my goal was to determine the relationship between *CBF* genes and LT tolerance. a) We determined that wheat has a large and complex *CBF* family; we identified and characterized 15 different *CBF* genes from hexaploid wheat b) Poaceae *CBF*s can be classified into 10 groups that share a common phylogenetic origin and similar structural characteristics. c) The wheat *CBF* family is representative in both size and complexity of *CBF* families in cereals; d) wheat *CBFs* display structure and functional characteristics of CBF factors; e) *TaCBF* shows differences in expression profile in winter wheat than in spring wheat, the difference affecting the LT tolerance and two groups *TaCBF* gene are candidates for Triticeae LT-tolerance. The results suggest that, as in dicots, CBFs are also an important component of the cold response pathway of cereals.

Two ICE genes were identified in wheat; *TaICE87* and *TaICE41*, encode a MYC-like bHLH transcriptional activator. TaICE87 and TaICE41 bind differently and specifically to the MYC recognition sequences in the *TaCBFIVd-B9* promoter. *TaICE87* and *TaICE41* are expressed constitutively at low levels, and their overexpressions in wild-type plants enhance the expression of the *CBF* regulon in the cold and improve freezing tolerance of the transgenic plants. These results show that the CBF/DREB regulon and regulation are evolutionarily conserved between monocot and dicot.

PERSPECTIVES

The higher inherited and inducible expressions of the five CBF specific groups in winter wheat suggest that these groups may play a predominant role in the winter cultivars superior LT tolerance capacity. It will be necessary to isolate all promoters of these *CBF* groups from spring and winter wheat, screen for specific trans-acting proteins regulating these *CBF* groups in winter wheat. Negative regulations as well as positive regulation are important for gene expression. The degradation of transcription factor proteins plays an important role in the negative regulation of gene expression. Since the control of mRNA stability is an important mode of gene regulation in response to environmental stimuli in plants, we will determine if CBF levels are due to a higher stability of the transcript in the winter cultivars or the instability in the spring ones due to specific microRNA that target the spring wheat mRNA for degradation.

Cold stresse affect plant growth and development, such as flowering time. Transcription is important in regulating plant development and environmental interactions, which may be affected by cross talk in transcriptional regulatory networks. Cross talk and link between signal transduction pathways will be an important subject in the future. The key vernalization gene VRN1 was mapped to the *Vrn-1* regions of homeologous group 5 chromosomes, regions that are associated with vernalization and FT in wheat .Genetic analyses have identified two loci in wheat and barley that mediate the capacity to overwinter. One locus is known as frost resistance-1(Fr-1) and co-segregates with the vernalization locus Tavrt1/VRN1. The second locus, Fr-2 lies in a region containing a cluster of more than 12 *CBF* genes. Our recent study on *CBF* genes in wheat demonstrates that five CBF groups mapped to group 5 chromosomes display higher constitutive and LT inducible expression in the winter cultivar. These evidences corroborate previous observations that plants which had reached flowering competence had a limited capacity to cold acclimation, and

support the hypothesis that the vernalization pathway exerts a regulatory role and is genetically linked to the pathway inducing the development of FT in cereals. Understanding this link may allow for a more effective manipulation of both FT and vernalization to produce cereal varieties with higher and more sustained tolerance to abiotic stress. The elucidation of the regulatory mechanisms that control these important traits in a complex system such as wheat will facilitate the development of the appropriate strategies to manipulate multigenic traits such as FT in agronomically-important crops. This will help in breeding new cultivars that adapt to current and future climate changes.

GENERAL REFERENCES

Agarwal, M., Hao, Y., Kapoor, A., Dong, C.-H., Fujii, H., Zheng, X., and Zhu, J.-K. (2006). A R2R3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. J Biol Chem 281: 37636–37645.

Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., and Suzuki, M. (1998). A novel mode of DNA recognition by a β-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. EMBO J 17: 5484-5496.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., et al (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science 301: 653–657.

Badawi, M., Danyluk, J., Boucho, B., Houde, M., and Sarhan, F. (2007). The *CBF* gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs. Mol Genet Genomics 277: 533-554.

Båga, M., Chodaparambil, S.V., Limin, A.E., Pecar, M., Fowler, D.B., Chibbar, R..N. (2006). Identification of quantitative trait loci and associated candidate genes for low-temperature tolerance in cold-hardy winter wheat. Funct Integr Genomics (in press). doi: 10.1007/s10142-006-0030-7.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol. 2: 28-36.

Baker, S.S., Wilhem, K.S. and 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.

Banno, H., Ikeda, Y., Niu, Q.W., and Chua, N.H. (2001). Overexpression of Arabidopsis ESR1 induces initiation of shoot regeneration. Plant Cell 13: 2609–2618.

Benedict, C., Geisler, M., Trygg, J., Huner, N., and Hurry, V. (2006). Consensus by democracy. Using meta-analyses of microarray and genomic data to model the cold acclimation signaling pathway in *Arabidopsis*. Plant Physiol 141: 1219–1232.

Bohnert, H.J., Su, H., and Shen, B. (1999). Molecular mechanisms of salinity tolerance. In Molecular Responses to Cold, Drought, Heat, and Salt Stress in Higher Plants (eds K. Shinozaki & K.Yamaguchi-Shinozaki), P. 29–62. R.G. Landes Company, Austin, TX.

Bowman, J.L., Alvarez, J., Weigel, D., Meyerowitz, E.M. and Smyth, D.R. (1993). Control of flower development in *Arabidopsis thaliana* by *APETALA1* and interacting genes. Development 119: 721–743.

Bräutigam, M., Lindlöf, A., Zakhrabekova, S., Gharti-Chhetri, G., Olsson, B., and Olsson, O. (2005). Generation and analysis of 9792 EST sequences from cold acclimated oat, *Avena sativa*. BMC Plant Biol 5:18.

Breton, G., Ouellet, F., Danyluk, J. and Sarhan, F. (2000). Biotechnological applications of plant freezing associated proteins. Biotechnology Annual Review.6: 57-100.

Brown, R.L., Kazan, K., McGrath, K.C, Maclean, D.J. and Manners, J.M. (2003). A role for the GCC-box in jasmonate-mediated activation of the PDF1.2 gene of *Arabidopsis*. Plant Physiol 132: 1020–1032.

Callebaut, I., Labesse, G., Durand, P., Poupon, A., Canard, L., Chomilier, J., Henrissat, B., and Mornon, J.P. (1997). Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. Cell Mol Life Sci 53: 621-645.

Carroll, S.B., Grenier, J.K. and Weatherbee, S.D. (2001). From DNA to Diversity: Molecular Genetics and the Evolution of Animal Design, Blackwell Scientific, Malden, MA.

Chakravarthy, S., Tuori, R.P., D'Ascenzo, M.D., Fobert, P.R., Despres, C., and Martin, G.B. (2003). The Tomato Transcription Factor Pti4 Regulates Defense-Related Gene Expression via GCC Box and Non-GCC Box cis Elements. Plant Cell 15: 3033-3050.

Chen, X. (2004). A microRNA as a translational repressor of APETALA2 in *Arabidopsis* flower development. Science 303: 2022-2025.

Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B.H., Hong, X., Agarwal, M., and Zhu, J.-K. (2003). ICE1: A regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. Genes Dev 17: 1043–1054.

Chinnusamy, V., Zhu, J., and Zhu, J.-K. (2006). Gene regulation during cold acclimation in plants. Physiol Plant 126: 52–61.

Choi, D.W., Rodriguez, E.M., and Close, T.J. (2002). Barley CBF3 gene identification, expression pattern, and map location. Plant Physiol 129: 1781–1787.

Clough, S.J. and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J 16: 735–743.

Cook, D, Fowler, S., Fiehn, O., and Thomashow, M.F. (2004). A prominent role for the CBF cold response pathway in configuring the low temperature metabolome of Arabidopsis. Proc Natl Acad Sci USA 101: 15243-15248.

Danyluk, J., Kane, N.A., Breton, G., Limin, A.E., Fowler, D.B. and Sarhan, F. (2003). TaVRT-1, a putative transcription factor associated with vegetative to reproductive transition in cereals. Plant Physiol 132: 1849–1860.

Danyluk, J, and Sarhan, F. (1990). Differential mRNA transcription during the induction of freezing tolerance in spring and winter wheat. Plant Cell Physiol 31: 609-619.

DeVries, A.L. (1971). Glycoproteins as biological antifreeze agents in antarctic fishes. Science 172: 1152-5.

Delauney, A. and Verma, D.P.S. (1993). Proline biosynthesis and osmoregulation in plants. Plant J 4: 215–223.

Dong, C.-H., Agarwal, M., Zhang, Y., Xie, Q., and Zhu, J.-K. (2006). The negative regulator of plant cold responses, HOS1, is a RING E3 ligase that mediates the ubiquitination and degradation of ICE1. Proc Natl Acad Sci USA 103: 8281–8286.

Drews, G.N., Bowman, J.L. and Meyerowitz E.M., (1991). Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the APETALA2 product. Cell 65: 991–1002.

Dubouzet, J.G., Sakuma Y., Ito Y., Kasuga M., Dubouzet E.G., Miura S., Seki M., Shinozaki K., Yamaguchi-Shinozaki K. (2003). OsDREB1A genes in rice, *Oryza sativa* L. encode transcription activators that function in drought, high-salt and cold-responsive gene expression. Plant J 33: 751-763.

Endo, T.R. (1988). Induction of chromosomal structural changes by a chromosome of *Aegilops cylindrica* L in common wheat. J Hered 79:366–370.

Endo, T.R. and Gill, B.S. (1996) The Deletion Stocks of Common Wheat. J Hered 87: 295-307.

Feldman, M. (2001). Origin of cultivated wheat. In The World Wheat Book, A History of Wheat Breeding. Edited by Bonjean, A.P. and Angus, W.J. pp. 3–56. Lavoisier Publishing, Londres.

Feng, J-X., Liu, D., Pan, Y., Gong, W., Ma, L-G., Luo, J-C., Deng, X.W., and Zhu, Y-X. (2005). An annotation update via cDNA sequence analysis and comprehensive profiling of developmental, hormonal or environmental responsiveness of the *Arabidopsis* AP2/EREBP transcription factor gene family. Plant Mol Biol 59: 853–868.

Fowler, S.G., Cook D., and Thomashow M.F. (2005). Low temperature induction of *Arabidopsis* CBF1, 2, and 3 is gated by the circadian clock. Plant Physiol 137: 961–968.

Francia, E., Rizza, F., Cattivelli, L., Stanca, A..M., Galiba, G., Tóth, B., Hayes, P.M., Skinner, J.S., and Pecchioni, N. (2004). Two loci on chromosome 5H determine low-temperature tolerance in a 'Nure' (winter) x 'Tremois' (spring) barley map. Theor Appl Genet 108: 670–680.

Fujimoto, S.Y., Ohta, M., Usui, A., Shinshi, H., and Ohme-Takagi, M. (2000). Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. Plant Cell 12: 393-404.

Furihata, T., Maruyama, K., Fujita, Y., Umezawa, T., Yoshida, R., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2006). Abscisic acid-dependent multisite phosphorylation regulates the activity of a transcription activator AREB1. Proc. Natl. Acad. Sci. USA 103:1988–1993.

Gao, M.J., Allard, G., Byass, L., Flanagan, A.M., and Singh, J. (2002). Regulation and characterization of four CBF transcription factors from *Brassica napus*. Plant Mol Biol 49: 459–471.

Gaboriaud, C., Bissery, V., Benchetrit, T., and Mornon, J.P. (1987). Hydrophobic cluster analysis: an efficient new way to compare and analyse amino acid sequences. FEBS Letters 224: 149-155.

Gill, B.S., Appels, R., Botha-Oberholster, A-M., Buell, C.R., Bennetzen, J.L., Chalhoub, B., Chumley, F., Dvořák, J., Iwanaga, M., Keller, B., Li, W., McCombie, W.R., Ogihara, Y., Quetier, F., and Sasaki, T. (2004). A workshop report on wheat genome sequencing: International Genome Research on Wheat Consortium. Genetics 168: 1087–1096.

Gilmour, S.J., Fowler, S.G., and Thomashow, M.F. (2004). *Arabidopsis* transcriptional activators CBF1, CBF2, and CBF3 have matching functional activities. Plant Mol Biol 54: 767-781.

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

Gilmour, S.J., Zarke, D.G., Stockinger, E.J., Salazar, M.P., Houghton, J.M., and Thomashow, M.F. (1998). Low temperature regulation of the *Arabidopsis* CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression. Plant J 16: 433-442.

Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H., Hadley, D., Hutchison, D., Martin, C., Katagiri, F., Lange, B.M., Moughamer, T., Xia, Y., Budworth, P., Zhong, J., Miguel, T., Paszkowski, U., Zhang, S., Colbert, M., Sun, W.L., Chen, L., Cooper, B., Park, S., Wood, T.C., Mao, L., Quail, P., Wing, R., Dean, R., Yu, Y., Zharkikh, A., Shen, R., Sahasrabudhe, S., Thomas, A., Cannings, R., Gutin, A., Pruss, D., Reid, J., Tavtigian, S., Mitchell, J., Eldredge, G., Scholl, T., Miller, R.M., Bhatnagar, S., Adey, N., Rubano, T., Tusneem, N., Robinson, R., Feldhaus, J., Macalma, T., Oliphant, A., and Briggs, S. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). Science 296: 92-100.

Gong, Z., Lee, H., Xiong, L., Jagendorf, A., Stevenson, B., and Zhu, J.K. (2002). RNA helicase-like protein as an early regulator of transcription factors for plant chilling and freezing tolerance. Proc Natl Acad Sci USA 99: 11507–11512.

Gong, Z., Dong, C.H., Lee, H., Zhu, J., Xiong, L., Gong, D., Stevenson, B., and Zhu, J.K. (2005). A DEAD box RNA helicase is essential for mRNA export and important for development and stress responses in *Arabidopsis*. Plant Cell 17: 256–267.

Gutterson, N., and Reuber, T.L. (2004). Regulation of disease resistance pathways by AP2/ERF transcription factors. Curr Opin Plant Biol 7: 465–471.

Guy, C.L. (1990). Cold acclimation and freezing stress tolerance: role of protein metabolism. Annu Rev Plant Physiol Plant Mol Biol 41: 187–223.

Haake, V., Cook, D., Riechmann, J.L., Pineda, O., Thomashow, M.F., and Zhang, J.Z. (2002). Transcription factor CBF4 is a regulator of drought adaptation in *Arabidopsis*. Plant Physiol 130: 639-48.

Hao, D., Ohme-Takagi, M., and Sarai, A. (1998). Unique mode of GCC box recognition by the DNA-binding domain of ethylene responsive element-binding factor (ERF domain) in plants. J. Biol.Chem. 273: 26857–26861.

Hannah, M.A., Wiese, D., Freund, S., Fiehn, O., Heyer, A.G., and Hincha, D.K. (2006). Natural genetic variation of freezing tolerance in *Arabidopsis*. Plant Physiol 142: 98–112.

Harlan, J. R. (1995). The living fields: our agricultural heritage. Cambridge.

Hayes, J.D. and Pulford, D.J. (1995). The glutathione S-transferase supergene family. Regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance. Crit Rev Biochem Mol Biol 30: 445–460.

Heim, M. A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B., and Bailey, P. C. (2003). The basic helix-loop-helix transcription factor family in plants: A genome-wide study of protein structure and functional diversity. Mol Biol Evol 20: 735-747.

Higo, K., Ugawa, Y., Iwamoto, M., Korenaga, T. (1999). Plant cis-acting regulatory DNA elements (PLACE) database. Nucleic Acids Research 27: 297–300.

Houde, M., Dhindsa, R.S, Sarhan F. (1992). A molecular marker to select for freezing tolerance in Gramineae. Mol Gen Genet 234:43–48.

Houde, M., Belcaid, M., Ouellet, F., Danyluk, J., Monroy, A.F., Dryanova, A., Gulick, P., Bergeron, A., Laroche, A., Links, M.G., MacCarthy, L., Crosby, W.L., and Sarhan, F. (2006). Wheat EST resources for functional genomics of abiotic stress. BMC Genomics 7: 149.

Hsieh, T.H., Lee, J.T., Yang, P.T., Chiu, L.H., Charng, Y.Y., Wang, Y.C., and Chan, M.T. (2002a). Heteroloy expression of the *Arabidopsis* C-Repeat, dehydration response element binding factor1 genes confers elevated tolerance to chilling and oxidative stresses in transgenic tomato. Plant Physiol 129: 1-9.

Hsieh, T.H., Lee, J.T., Charng, Y.Y., and Chan, M.T. (2002b). Tomato plants ectopically expressing *Arabidopsis* CBF1 show enhanced resistance to water deficit stress. Plant Physiol 130: 618-626.

Hu, Y.X., Wang, Y.X., Liu, X.F., and Li, J.Y. (2004). *Arabidopsis* RAV1 is down-regulated by brassinosteroid and may act as a negative regulator during plant development. Cell Res 14: 8–15.

Huang, T., Nicodemus, J., Zarka, D.G., Thomashow, M.F., Wisniewski, M., and Duman, J.G. (2002a). Expression of an insect (*Dendroides canadensis*) antifreeze protein in *Arabidopsis thaliana* results in a decrease in plant freezing temperature. Plant Mol Biol 50: 333-44.

Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R., and Gornicki, P (2002b). Phylogenetic analysis of the acetyl-CoA carboxylase and 3-phosphoglycerate kinase loci in wheat and other grasses. Plant Mol Biol 48: 805–820.

Huang, S., Sirikhachornkit, A., Su, X., Faris, J., Gill, B., Haselkorn, R., and Gornicki, P. (2002c). Genes encoding plastid acetyl-CoA carboxylase and 3-phosphoglycerate kinase of the Triticum/Aegilops complex and the evolutionary history of polyploid wheat. Proc Natl Acad Sci USA 99: 8133–8138.

Huang, X., and Madan, A. (1999). CAP3: A DNA sequence assembly program. Genome Res 9: 868-877.

Ishitani, M., Xiong L., Lee, H., Stevenson, B., and Zhu, J.K. (1998). HOS1, a genetic locus involved in cold-responsive gene expression in Arabidopsis. Plant Cell 10: 1151–1161.

Ito, Y., Katsura K., Maruyama K., Taji T., Kobayashi M., Seki M., Shinozaki K., and Yamaguchi-Shinozaki K. (2006). Functional analysis of rice DREB1/CBF-type transcription factors involved in cold-responsive gene expression in transgenic rice. Plant Cell Physiol 47: 141–153.

Jaglo, K.R., Kleff, S., Amundsen, K.L. Zhang, X., Haake, V., Zang, J.Z., Deits, T., and Thomashow, M.F. (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 Physiol 127: 910-917.

Jaglo-Ottosen, K.R., Glimour, S.J., Zarka, D.G., and Thomashow, M.F. (1998). Arabidopsis CBF1 overexpression induced COR genes and enhances freezing tolerance. Science 280: 104-106.

Jiang, C., Iu, B., and Singh, J. (1996). Requirement of a CCGAC *cis*-acting element for cold induction of the *BN115* gene from winter *Brassica napus*. Plant Mol Biol. 30: 679–684.

Jofuku, K.D., denBoer, B.G., Van Montagu, M., and Okamuro, J.H.K. (1994). Control of *Arabidopsis* flower and seed development by the homeotic gene APETALA2. Plant Cell 6: 1211-1225.

Kagaya, Y., Ohmiya, K., and Hattori, T. (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 Res 27: 470-478.

Kang, J.Y., Choi, H.I., Im, M.Y. and Kim, S.Y. (2002). *Arabidopsis* basic leucine zipper proteins that mediate stress-responsive abscisic acid signaling, Plant Cell 14: 343–357.

Kasuga, M., Liu, Q., Miura, S., Yamaguchi-Shinozaki, K. and Shinozaki, K. (1999). Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor, Nature Biotechnol. 17: 287–291.

Keck, E., McSteen, P., Carpenter, R., and Coen, E. (2003). Separation of genetic functions controlling organ identity in flowers. EMBO J 5: 1058-66.

Kellogg, E.A. (2001). Evolutionary history of the grasses. Plant Physiol 125: 1198–1205.

Kim, S., Kang, J.Y., Cho, D.I., Park, J.H. and Kim, S.Y. (2004). ABF2, an ABRE-binding bZIP factor, is an essential component of glucose signaling and its overexpression affects multiple stress tolerance. Plant J 40: 75–87.

Klucher, K.M., Chow, H., Reiser, L. and Fischer, R.L. (1996). The *AINTEGUMENTA* gene of *Arabidopsis* required for ovule and female gametophyte development is related to the floral homeotic gene *APETALA2*. Plant Cell 8: 137–153.

Knight, M.R. (2002). Signal transduction leading to low-temperature tolerance in *Arabidopsis thaliana*. Philos Trans R Soc Lond B Biol Sci 357: 871–875.

Knight, H., Veale, E.L., Warren, G.J., Knight, M.R. (1999). The sfr6 mutation in Arabidopsis suppresses low-temperature induction of genes dependent on the CRT/DRE sequence motif. Plant Cell 11: 875–886.

Koch, M.A., Haubold, B., and Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, *Arabis*, and related genera (Brassicaceae). Mol Biol Evol 17: 1483-1498.

Kreps, J.A., Wu, Y., Chang, H.S., Zhu, T., Wang, X. and Harper, J.F. (2002). Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. Plant Physiol 130: 2129-41.

Kumar, S., Tamura K., and Nei M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5: 150-163.

Kume, S., Kobayashi, F., Ishibashi, M., Ohno, R., Nakamura, C., and Takumi, S. (2005). Differential and coordinated expression of Cbf and Cor/Lea genes during long-term cold acclimation in two wheat cultivars showing distinct levels of freezing tolerance. Genes Genet Syst 80: 185-197.

Langridge, P., Paltridge, N., and Fincher, G. (2006). Functional genomics of abiotic stress tolerance in cereals. Briefings in Functional Genomics and Proteomics 4: 343-354.

Lee, B.-H., Henderson, D.A., and Zhu, J.-K. (2005). The *Arabidopsis* cold-responsive transcriptome and its regulation by ICE1. Plant Cell 17: 3155–3175.

Lee, B.J., Park, C.J., Kim, S.K., Kim, K.J. and Paek, K.H. (2006). In vivo binding of hot pepper bZIP transcription factor CabZIP1 to the G-box region of pathogenesis-related protein 1 promoter, Biochem. Biophys. Res. Comm. 344: 55–62.

Lee, H., Xiong, L., Gong, Z., Ishitani, M., Stevenson, B., and Zhu, J.K. (2001). The Arabidopsis HOS1 gene negatively regulates cold signal transduction and encodes a RING finger protein that displays cold-regulated nucleo-cytoplasmic partitioning. Genes Dev 15: 912–924.

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P., and Rombauts, S. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. Nucleic Acids Res. 30: 325–327.

Li, J., Brader, G., and Palva, E.T. (2004). The WRKY70 transcription factor: a node of convergence for jasmonate-mediated and salicylate-mediated signals in plant defense. Plant Cell. 16: 319-331.

Limin, A.E., Danyluk, J., Chauvin, L.P., Fowler, D.B., and Sarhan, F. (1997). Chromosome mapping of low-temperature induced WCS120 family genes and regulation of cold-tolerance expression in wheat. Mol Gen Genet 253: 720-727.

Liu, Q., Kasuga, M., Sakuma, Y., Abe, H., Miura, S., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1998). Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal pathways in drought- and low temperature- responsive gene expression, respectively, in *Arabidopsis*. Plant Cell 10: 1391-1406.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCt method. Methods 25: 402–408.

Magome, H., Yamaguchi, S., Hanada, A., Kamiya, Y., and Oda, K. (2004). Dwarf and delayed-flowering 1, a novel *Arabidopsis* mutant deficient in gibberellin biosynthesis because of overexpression of a putative AP2 transcription factor. Plant J 37: 720-729.

Magnani, E., Sjölander, K., and Hake, S. (2004). From endonuclease to transcription factors: evolution of the AP2 DNA binding domain in plants. Plant Cell 16: 2265–2277.

Marchler-Bauer, A., Anderson, J.B., Cherukuri, P.F., DeWeese-Scott, C., Geer, L.Y., Gwadz, M., He,S., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Liebert, C.A.,

Liu, C., Lu, F., Marchler, G.H., Mullokandov, M., Shoemaker, B.A., Simonyan, V., Song, J.S., Thiessen, P.A., Yamashita, R.A., Yin, J.J., Zhang, D., and Bryant, S.H. (2005). CDD: a Conserved Domain Database for protein classification. Nucleic Acids Res. 33: 192-196.

Medina, J., Bargues, M., Terol, J., Perez-Alonso, M., and 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 Physiol 119: 463–470.

Miller, A.K., Galiba, G., and Dubcovsky, J. (2006). A cluster of 11 CBF transcription factors is located at the frost tolerance locus Fr-Am2 in *Triticum monococcum*. Mol Genet Genomics 275: 193-203.

Minty, A., Dumont, X., Kaghad, M., and Caput, D. (2000). Covalent modification of p73 α by SUMO-1. Two-hybrid screening with p73 identifies novel SUMO-1-interacting proteins and a SUMO-1 interaction motif. J Biol Chem 275: 36316–36323.

Miura, K., Jin, J.O., Lee, J., Yoo, C.Y., Stirm, V., Miura, T., Ashworth, E.N., Bressan R.A., Yun, Dae-J. and Hasegawa P.M. (2007). SIZ1-Mediated Sumoylation of ICE1 Controls CBF3/DREB1A Expression and Freezing Tolerance in Arabidopsis. Plant Cell 19: 1403-1414.

Mizukami, Y., (2001). A matter of size: developmental control of organ size in plants, Curr. Opin. Plant Biol. 4: 533–539.

Nakagoshi, H., Nagase, T., Kanei-Ishii, C., Ueno, Y., and Ishii, S. (1990). Binding of the c-myb proto-oncogene product to the simian virus 40 enhancer stimulates transcription. J Biol Chem. 265: 3479–3483.

Nakano, T., Suzuki, K., Fujimura, T., and Shinshi, H. (2006). Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. Plant Physiol 140: 411–432.

NDong, C., Danyluk, J., Wilson, K.E., Pocock, T., Huner, N.P., and Sarhan, F. (2002). Cold-regulated cereal chloroplast late embryogenesis abundant-like proteins. Molecular characterization and functional analyses. Plant Physiol. 129: 1368-81.

Novillo, F., Alonso, J.M., Ecker, J.R., and 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*. Proc Natl Acad Sci USA 101: 3985-3990.

Oh S.J., Song, S.I.K., Kim, Y.S., Jang, H.J., Kim, S.Y., Kim, M., Kim, Y.K., Nahm, B.H., Kim, J.K. (2005). Arabidopsis CBF3/DREB1A and ABF3 in transgenic rice increased tolerance to abiotic stress without stunting growth. Plant Physiol 138: 341–351.

Ohto, M.A., Fischer, R.L., Goldberg, R.B., Nakamura, K., and Harada, J.J. (2005). Control of seed mass by APETALA2. Proceedings of the National Academy of Sciences, USA 102: 3123–3128.

Ohme-Takagi, M., and Shinshi, H. (1995). Ethylene-inducible DNA binding proteins that interact with an ethylene-responsive element. Plant Cell 7: 173–182.

Ouellet, F., Vazquez-Tello, A., and Sarhan, F. (1998). The wheat *wcs120* promoter is cold-inducible in both monocotyledonous and dicotyledonous species. FEBS Lett. 423: 324–328.

Papp, B., Pál, C., and Hurst, L.D. (2004). Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. Nature 429: 661-664.

Parenicova, L., de Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., Angenent, G.C., and Colombo, L. (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. Plant Cell 15: 1538-1551.

Park, J.M., Park, C.J., Lee, S.B., Ham, B.K., Shin, R. and Paek, K.H. (2001). Overexpression of the tobacco Tsi1 gene encoding an EREBP/AP2-type transcription factor enhances resistance against pathogen attack and osmotic stress in tobacco. Plant Cell 13: 1035–1046.

Ptashne, M. (1988). How eukaryotic transcriptional activators work. Nature 335: 683–689.

Qin, F., Sakuma, Y., Li, J., Liu, Q., Li, Y-Q., Shinozaki, K., and Yamaguchi-Shinozaki, K. (2004). Cloning and functional analysis of a novel DREB1/CBF transcription factor involved in cold-responsive gene expression in *Zea mays* L. Plant Cell Physiol 45: 1042–1052.

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005). InterProScan: protein domains identifier. Nucleic Acids Res 33: W116-W120.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C.Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., et al (2000). *Arabidopsis* transcription

factors: genome-wide comparative analysis among eukaryotes. Science 290: 2105–2110.

Riechmann, J.L., and Meyerowitz, E.M. (1998). The AP2/EREBP family of plant transcription factors. Bio Chem 379: 633-646.

Rozen, S., and Skaletsky H.J. (2000). Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, pp 365-386.

Sakai, A., and Larcher, W. (1987). Frost survival of plants: Responses and adaptation to freezing stress. Springer-Verlag, Berlin.

Sakuma, Y., Liu, Q., Doubozet, J.G., Abe, H., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2002). DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration and cold-induced gene expression. Biochem Biophys Res Commun 290: 998-1009.

Sears, E.R. (1954). The Aneuploids of Common Wheat. (Columbia, MO.: University of Missouri Agricultural Experiment Station,), pp. 1-58.

Sears, E.R. (1966). Nullisomic-tetrasomic combinations in hexaploid wheat. Oliver & Boyd, ed (Edinburgh: R. Riley and K. R. Lewis), pp. 29-45.

Seki, M., Narusaka, M., Abe, H., Kasuga, M., Yamaguchi-Shinozaki, K., Carninci, P., Hayashizaki, Y., Shinozaki, K. (2001). Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray. Plant Cell 13: 61–72.

Seki, M., Narusaka, M., Ishida, J., Nanjo, T., Fujita, M., Oono, Y., Kamiya, A., Nakajima, M., Enju, A., Sakurai, T., Satou, M., Akiyama, K., Taji, T., Yamaguchi-Shinozaki, K., Carninci, P., Kawai, J., Hayashizaki, Y., and Shinozaki, K. (2002). Monitoring the expression profiles of 7000 *Arabidopsis* genes under drought, cold and high-salinity stresses using a full-length cDNA microarray. Plant J 31: 279–292.

Shigyo, M. and Ito, M. (2004). Analysis of gymnosperm two-AP2-domain-containing genes. Dev. Genes Evol. 214: 105–114.

Shigyo, M., Hasebe, M. and Ito, M. (2006). Molecular evolution of the AP2 subfamily. Gene 366: 256-265.

Shinozaki, K., Yamaguchi-Shinozaki, K., and Seki, M. (2003). Regulatory network of gene expression in the drought and cold stress responses. Curr Opini in Plant Biol 6: 410-417.

Shinozaki, K. and Yamaguchi-Shinozaki, K. (2006). Global Analysis of Networks to Solve Complex Abiotic Stress Responses in Cold Hardiness in Plants Molecular Genetic, Cell Biology and Physiology Editors: T.H.H. Chen, M. Uemura and S. Fujikawa. CAB International. Pp.1-10.

Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000). Molecular responses to dehydration and low temperature: differences and cross-talk between two stress signaling pathways. Curr. Opin. Plant Biol. 3: 217-223.

Shinozaki, K. and Yamaguchi-Shinozaki, K. (1997). Gene expression and signal transduction in water-stress response. Plant Physiol. 115: 327-334.

Shinwari, Z.K., Nakashima, K., Miura, S., Kasuga, M., Seki, M., Yamaguchi-Shinozaki, K., and Shinozaki, K. (1998). An *Arabidopsis* gene family encoding DRE/CRT binding proteins involved in low-temperature-responsive gene expression. Biochem Biophys Res Commun 250: 161–170.

Skinner, J.S., von Zitzewitz, J., Szucs, P., Marquez-Cedillo, L., Filichkin, T., Amundsen, K., Stockinger, E.J., Thomashow, M.F., Chen, T.H., and Hayes, P.M. (2005). Structural, functional, and phylogenetic characterization of a large CBF gene family in barley. Plant Mol Biol 59: 533-551.

Skinner, J.S., Szűcs, P., von Zitzewitz J., Marquez-Cedillo, L., Filichkin, T., Stockinger, E.J., Thomashow, M.F., Chen, T.H.H., and Hayes, P.M. (2006). Mapping of barley homologs to genes that regulate low temperature tolerance in *Arabidopsis*. Theor Appl Genet 112: 832–842.

Sohn, K.H., Lee, S.C., Jung, H.W., Hong, J.K., and Hwang, B.K. (2006). Expression and functional roles of the pepper pathogen-induced transcription factor RAV1 in bacterial disease resistance, and drought and salt stress tolerance. Plant Mol Biol 61: 897–915.

Stepanova, A.N. and Ecker, J.R. (2000). Ethylene signaling: From mutant to molecules. Curr. Opin. Plant Biol.3: 353-360.

Stockinger, E.J., Glimour, S.J. and Thomashow, N.F. (1997). *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that bind to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. Proc Natl Acad Sci USA 94: 1035-1040.

Suzuki, M., Kao, C.Y., and McCarty, D.R. (1997). The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity. Plant Cell 9: 799–807.

Swofford, D.L. (2003) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4. (Sunderland, MA: Sinauer Associates).

Takumi, S., Koike, A., Nakata, M., Kume, S., Ohno, R. and Nakamura, C. (2003). Cold-specific and light-stimulated expression of a wheat (*Triticum aestivum* L.) Cor gene Wcor15 encoding a chloroplast-targeted protein. J Exp Bot. 391: 2265-74.

Tang, M., Lü, S., Jing, Y., Zhou, X., Sun, J., and Shen, S. (2005). Isolation and identification of a cold-inducible gene encoding a putative DRE-binding transcription factor from Festuca arundinacea. Plant Physiol Biochem 43: 233–239.

Thomashow, M.F. (1999). Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. Ann Rev Plant Physiol Plant Mol Biol 50: 571-599.

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

Tilman, D., Cassman, K.G., Matson, P.A., Naylor, R., Polasky, S. (2002). Agricultural sustainability and intensive production practices. Nature 418: 671-677.

Toledo-Ortiz, G., Huq, E., and Quail, P.H. (2003). The *Arabidopsis* basic/helix–loop–helix transcription factor family. Plant Cell 15: 1749–1770.

Tondelli, A., Francia, E., Barabaschi, D., Aprile, A., Skinner, J. S., Stockinger, E. J., Stanca, A. M., and Pecchioni, N. (2006). Mapping regulatory genes as candidates for cold and drought stress tolerance in barley. Theor. Appl. Genet. 112: 445-454.

Uno, Y., Furihata, T., Abe, H., Yoshida, R., Shinozaki, K. and Yamaguchi-Shinozaki, K. (2000). *Arabidopsis* basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions.Proc. Natl. Acad. Sci. U. S. A. 97: 11632–11637.

Vágújfalvi, A., Aprile A., Miller A., Dubcovsky J., Delugu G., Galiba G., and Cattivelli L. (2005). The expression of several Cbf genes at the Fr-A2 locus is linked to frost resistance in wheat. Mol Genet Genomics 274: 506-514.

Van Buskirk, H.A., and Thomashow, M.F. (2006). *Arabidopsis* transcription factors regulating cold acclimation. Physiol Plant 126: 72–80.

Vierstra, R.D. (1996). Proteolysis in plants: mechanisms and functions. Plant Mol Biol 32: 275-302.

Vogel, J.T., Zarka, D.G., Van Buskirk, H.A., Fowler, S.G., and Thomashow, M.F. (2005). Roles of the CBF2 and ZAT12 transcription factors in configuring the low temperature transcriptome of *Arabidopsis*. Plant J 41: 195–211.

Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. Plant J 33: 949-956.

Wang, H., Huang, Z.J., Chen, Q. Zhang, Z., Zhang, H., Wu, Y. , Huang, D. and Huang, R. (2004). Ectopic overexpression of tomato JERF3 in tobacco activates downstream gene expression and enhances salt tolerance. Plant Mol Biol 55: 183–192.

Wang, Z., Triezenberg, S.J., Thomashow, M.F., and Stockinger, E.J. (2005). Multiple hydrophobic motifs in *Arabidopsis* CBF1 COOH-terminus provide functional redundancy in trans-activation. Plant Mol Biol 58: 543–559.

Wei, G., Pan, Y., Lei, J., and Zhu, Y-X. (2005). Molecular cloning, phylogenetic analysis, expressional profiling and in vitro studies of TINY2 from *Arabidopsis* thaliana. J Biochem Mol Biol 38: 440-446.

Weretilnyk, E.A., and Hanson, A.D. (1989). Betaine aldehyde dehydrogenase from spinach leaves: purification, in vitro translation of the mRNA, and regulation by salinity. Arch Biochem Biophys 271: 56–63.

Wessler, S.R. (2005). Homing into the origin of the AP2 DNA binding domain. Trends Plant Sci 10: 54–56.

Wuitschick, J.D., Lindstrom, P.R., Meyer A.E., and Karrer, K.M. (2004). Homing endonuclease encoded by germ line-limited genes in *Tetrahymena thermophila* have APETALA2 DNA binding domains. Eukaryot Cell 3: 685–694.

Xin, Z., and Browse, J. (1998). Eskimo1 mutants of *Arabidopsis* are constitutively freezing-tolerant. Proc Natl Acad Sci USA 95: 7799–7804.

Xiong, Y, and Fei S-Z (2006). Functional and phylogenetic analysis of a DREB/CBF-like gene in perennial ryegrass (*Lolium perenne* L.). Planta 224: 878-888.

Xu, D., Duan X., Wang B., Hong B., Ho T.H., Wu R. (1996). Expression of a late embryogenesis abundant protein gene, *HVA1*, from barley confers tolerance to water deficit and salt stress in transgenic rice. Plant Physiol. 110: 249–257.

Xue, G-P (2002). Characterization of the DNA-binding profile of HvCBF1 using an enzymatic method for rapid, quantitative and high-throughput analysis of the DNA binding activity. Nucleic Acids Res 30: e77

Xue, G.P. (2003). The DNA-binding activity of an AP2 transcriptional activator HvCBF2 involved in regulation of low-temperature responsive genes in barley is modulated by temperature. Plant J 33: 373-83.

Xue, G.P. (2005). A CELD-fusion method for rapid determination of the DNA-binding sequence specificity of novel plant DNA-binding proteins. Plant J 41: 638–649.

Yamaguchi-Shinozaki, K., and Shinozaki, K. (1994). A novel cis-acting element in an *Arabidopsis thaliana* gene is involved in responsiveness to drought, low - temperature, or high-salt stress. Plant Cell 6: 251-264.

Yancey, P.H. (1994). Compatible and counteracting solutes. In K Strange, eds, Cellular and Molecular Physiology of Cell Volume Regulation. CRC Press, Boca Raton, FL, pp 81-109.

Zarka, D.G., Vogel, J.T., Cook, D., Thomashow, M.F. (2003). Cold induction of *Arabidopsis CBF* genes involves multiple ICE (inducer of CBF expression) promoter elements and a cold-regulatory circuit that is desensitized by low temperature. Plant Physiol 133: 910–918.

Zhao, T-J., Sun, S., Liu, Y., Liu, J-M., Liu, Q., Yan, Y-B., Zhou, H-M. (2006). Regulating the drought-responsive element (DRE)-mediated signaling pathway by synergic functions of trans-active and trans-inactive DRE binding factors in Brassica napus. J Biol Chem 281: 10752–10759.

Zhu, J-K. (2002). Salt and drought stress signal transduction in plants. Annual Review of Plant Biology 53: 247–273.

Zhu. J., Shi, H., Lee, B.H., Damsz, B., Cheng, S., Stirm, V., Zhu, J.K., Hasegawa, P.M., and Bressan, R.A. (2004). An *Arabidopsis* homeodomain transcription factor gene, HOS9, mediates cold tolerance through a CBF-independent pathway. Proc Natl Acad Sci USA 101: 9873–9878.

Zhu, J., Verslues, P.E., Zheng, X., Lee, B.H., Zhan, X., Manabe, Y., Sokolchik, I., Zhu, Y., Dong, C.H., Zhu, J.K., Hasegawa, P.M., and Bressan, R.A. (2005). HOS10 encodes an R2R3-type MYB transcription factor essential for cold acclimation in plants. Proc Natl Acad Sci USA 102: 9966–9971.