De Novo Regulatory Motif Discovery Identifies Significant Motifs in Promoters of Five Classes of Plant Dehydrin Genes

Yevgen Zolotarov, Martina Strömvik*

Department of Plant Science, Macdonald Campus, McGill University, 21111 Lakeshore Road, Sainte-Anne-de-Bellevue, QC, H9X 3V9, Canada

* martina.stromvik@mcgill.ca

Abstract

Plants accumulate dehydrins in response to osmotic stresses. Dehydrins are divided into five different classes, which are thought to be regulated in different manners. To better understand differences in transcriptional regulation of the five dehydrin classes, *de novo* motif discovery was performed on 350 dehydrin promoter sequences from a total of 51 plant genomes. Overrepresented motifs were identified in the promoters of five dehydrin classes. The K_n dehydrin promoters contain motifs linked with meristem specific expression, as well as motifs linked with cold/dehydration and abscisic acid response. KS dehydrin promoters contain a motif with a `GATA` core. SK_n and Y_nSK_n dehydrin promoters contain motifs that match elements connected with cold/dehydration, abscisic acid and light response. Y_nK_n dehydrin promoters contain motifs that match abscisic acid and light response elements, but not cold/dehydration response elements. Conserved promoter motifs are present in the dehydrin classes and across different plant lineages, indicating that dehydrin gene regulation is likely also conserved.

Introduction

Plants have developed specific mechanisms that allow them to prepare for and survive drastic changes in their environment. One of the better-studied mechanisms is cold acclimation, which allows plants to develop freezing tolerance [1,2]. During exposure to low non-freezing temperatures gene expression is modulated and numerous solutes, known as osmoprotectants, and protective proteins accumulate in plant tissues. Dehydrins or dehydration proteins, (DHN) belong to group II LEA (late embryogenesis abundant) proteins. They are often found among those protective proteins and they are ubiquitous in transcriptomes of plants under osmotic stress, such as cold, drought and high salinity [3–7]. All dehydrins contain a 15 amino acid K-segment, rich in lysine residues, represented by `EKKGIMDKIKEKLPG` conserved sequence [8]. The K-segment forms an amphipathic α-helix that allows dehydrins to stabilize plant membranes and proteins during dehydration stresses [9–12]. In addition to the K-segment, dehydrins can contain a Y-segment (T/VDEYGNP) and an S-segment (3+ serines)

[13]. The S-segment is thought to be involved in ion binding and dehydrin phosphorylation, which induces a conformational change in dehydrins [14,15]. Currently, the function of the Y-segment is unknown. The dehydrins are categorized into 5 subclasses (K_n, KS, SK_n, Y_nK_n and Y_nSK_n) based on the presence and location of the 3 conserved segments [13]. Members of each subclass are expressed in response to a different set of stimuli. However, there is no clear link between subclass types and expression triggers [16].

As defined by Close, 1997, dehydrins must contain a K-segment. According to that definition, dehydrins are only found in plants. There are other proteins described as dehydrins, for example from *Escherichia coli* (GenBank: AAB18249.1), a fungus *Pneumocystis carinii* (GenBank: CAC43457.1) [4] or from whitish truffle *Tuber borchii* (GenBank: ABC33908.1) [17]. However, these proteins do not contain the K-segment or any of the other conserved dehydrin segments and therefore should not be considered to be proper dehydrins.

Stress response in plants can be regulated in an abscisic acid (ABA) dependent and/or independent manner [18]. Multiple transcription factors, such as C-repeat binding factor/dehydration responsive element binding protein (CBF/DREB) and ABA response element binding protein (AREB), participate in water stress response, by binding to *cis*-regulatory elements in the promoters of their respective regulons. The CBF1-3 are transcription factors that participate in ABA independent cold and dehydration induced gene expression [19] and they bind a C-repeat (CRT) *cis*-regulatory element core (CCGAC), also known as dehydration response element (DRE). Members of the CBF regulon include well-studied *A. thaliana* genes, such as LTI78/COR78 [20], COR15A and COR47 (an SK_n dehydrin) [21]. However, not all members of the CBF regulon have the CRT *cis*-regulatory element in their promoters [22], hence there are yet undiscovered motifs that are involved in cold and drought response. Numerous transcription factors participate in the ABA dependent stress response and they bind several *cis*-regulatory elements with a `TACGTG` core [23]. Many members of the CBF regulon are also upregulated in response to drought and ABA exposure, demonstrating a cross-talk between stress-induced pathways [24]. For example, in barley (*Hordeum vulgare* L.), a K_n dehydrin is strongly upregulated in response to cold, dehydration and ABA, and its promoter contains CRT and abscisic acid response elements (ABREs) *cis*-regulatory elements, whereas a barley SK_n dehydrin, whose promoter contains multiple CRTs and no ABREs is only weakly upregulated in response to ABA, but shows a significant upregulation in response to cold [3]. The expression of CBFs, and, in turn, their regulons, is modulated by photoperiod through phytochrome B and phytochrome-interacting factors [25,26].

In this study, we tested whether the different classes of dehydrin genes house specific and conserved *cis*-regulatory elements in their promoters that could contribute to gene characterization. *De novo* motif discovery, a computational approach to identify statistically overrepresented sequence motifs within a promoter sequence, was used to analyze a total of 350 dehydrin promoters. For each of the five dehydrin classes, statistically significant motifs were identified, and matched to experimentally validated *cis*-regulatory elements known from literature. Motifs linked to ABA-dependent and ABA-independent stress response pathways were detected in the promoters of dehydrin genes from various, distant plant lineages, which indicates that the stress response pathways regulating dehydrin expression are conserved.

Methods

Plant genomes used in the computational analyses

Permission to use data from genomes that are not published was obtained from members of sequencing consortia, where stated. In other cases, published data was used.

The following genome sequences were obtained from Phytozome v10 (http://phytozome.jgi.doe.gov) [27] using BioMart [28]: *Amborella trichopoda* [29], *Aquilegia coerulea* (*Aquilegia coerulea* Genome Sequencing Project, http://www.phytozome.net/, permission obtained from Dr. Scott Hodges), *Arabidopsis halleri* (*Arabidopsis halleri* v1.1, DOE-JGI, http://www.phytozome.net/ahalleri), *Arabidopsis lyrata* [30], *Arabidopsis thaliana* [31], *Boechera stricta* (*Boechera stricta* v1.2, DOE-JGI, http://www.phytozome.net/bstricta), *Brachypodium distachyon* [32], turnip mustard (*Brassica rapa* L.) [33,34], papaya (*Carica papaya*) [35], *Capsella grandiflora* and *Capsella rubella* [36], clementine (*Citrus clementina*) and, sweet orange (*Citrus sinensis*) [37], cucumber (*Cucumis sativus*) (permission obtained from Dr. Yiqun Weng), *Eucalyptus grandis* [38], *Eutrema salsugineum* (formerly *Thellungiella halophila*) [39], strawberry (*Fragaria vesca*) [40], soybean (*Glycine max*) [41], cotton (*Gossypuim raimondii*) [42], flax (*Linum usitatissimum*) [43], apple (*Malus domestica*) [44], cassava (*Manihot esculenta*) [45], barrel medic (*Medicago truncatula*) [46], monkey flower (*Mimulus guttatus*) [47], rice (*Oryza sativa*) [48], swtichgrass (*Panicum virgatum* v1.0, DOE-JGI, http://www.phytozome.net/pvirgatum), common bean (*Phaseolus vulgaris* L.) [49], moss (*Physcomitrella patens*) [50], peach (*Prunus persica*) [51], poplar (*Populus trichocarpa*) [52], castor bean (*Ricinus communis*) [53], foxtail millet (*Setaria italica*) [54], Shrub willow (*Salix purpurea* v1.0, DOE-JGI, http://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Spurpurea), tomato (*Solanum lycopersicum*) [55], potato (*Solanum tuberosum*) [56], greater duckweed (*Spirodela polyrhiza*) [57], cocoa (*Theobroma cacao*) [58], grape (*Vitis vinifera*) [59], maize (*Zea mays*) [60].

The following genomes were obtained from other sources: kiwifruit (*Actinidia chinensis*) [61], sugar beet (*Beta vulgaris*) [62]; pigeonpea (*Cajanus cajan*) [63] http://www.icrisat.org/gt-bt/iipg/genomedata.zip; pepper (*Capsicum annuum*) [64]; chickpea (*Cicer arietinum*) [65] http://www.icrisat.org/gt-bt/ICGGC/genomedata.zip; *Lotus japonicus* [66] ftp://ftp.kazusa.or.jp/pub/lotus/lotus_r2.5/; banana (*Musa acuminata*) [67]; *Oryza brachyantha* [68]; date palm (*Phoenix dactylifera*, Draft Sequence Version 3) [69] http://qatar-weill.cornell.edu/research/datepalmGenome/download.html; Norway spruce (*Picea abies*) [70]; loblolly pine (*Pinus taeda*) [71].

Identification of dehydrin genes

A custom solution was used to identify all dehydrin genes found in the plant genomes described above. Amino acid sequences of several known dehydrins were obtained from NCBI GenBank [72] and sequences of their K-segment were used to populate a seed FASTA file. A Python script was written that used Biopython [73] Motif module to scan all amino acid sequence for proteins containing a sequence similar to the K-segment, based on its position frequency matrix (PFM). After each round of search new K-segment sequences were added to the original FASTA file. The Y-segment sequence file was constructed in a similar manner using identified dehydrin protein sequences. Identified dehydrins were categorized based on the occurrence of conserved segments using either their PFMs (K- and Y-segments) or a regular expression that described a simpler S-segment. All identified dehydrins were divided into five categories: K_n, KS, SK_n, Y_nK_n, Y_nSK_n and 1000 bps upstream of the transcription start site (where data was available, otherwise upstream from the start site) were obtained from Phytozome BioMart or they were directly extracted from the genomes using custom scripts. *Oxytropis arctobia* and *Oxytropis splendens* KS dehydrin gene sequences were obtained from NCBI GenBank (accessions: AEV59613 and AEV59617, respectively [6]). 1000 bp of *O. arctobia* and *O. splendens* promoters were obtained by amplifying GenomeWalker libraries and sequencing PCR products (Zolotarov et. al., unpublished).

To validate that the identified genes can actually be considered dehydrins, phmmer, as implemented on the HMMER web server [74] was used to search sequences on UniProt

Knowledgebase [75] that have a significant similarity to putative dehydrins discovered using our custom method. The top ten significant hits were taken for each putative dehydrin and their domain annotation was extracted. Additionally, Needleman-Wunsch [76] pairwise alignment was used to compare 15 putative KS dehydrin to a known *Arabidopsis* KS dehydrin (AT1G54410). The closest match to every putative dehydrin in the NCBI GenBank non-redundant database was searched for using BLAST [77].

All the scripts and sequence data used in this paper are available from https://github.com/zolotarov/dehydrin_promoters

Intrinsic disorder and hydrophilicity analysis

The identified dehydrin sequences were compared for intrinsic disorder and hydrophilicity with random plant protein sequences to assess the classification as a dehydrin. To calculate the grand average of hydrophilicity, Biopython ProtParam module was used. To calculate disorder proportion, IUPred [78] scores were calculated for each amino acid. The proportion of amino acids with the score above 0.5 (indicating disorder) was calculated. Statistical comparison was performed using t-test implemented in the scipy library [79,80]. The same number of random protein sequences was obtained for each species as the number of dehydrins used in this study. The sequences were downloaded using NCBI Entrez Direct E-utilities [81].

De novo motif discovery

Motifs were discovered using MEME v4.9.1 [82], Seeder v0.01 [83] and Weeder v1.4.2 [84], and using the five sets of sequences as separate input. Significant motifs were selected based on following parameters: E-value ≤ 0.05 for MEME, Q-value ≤ 0.01 for Seeder and the top 3 motifs recommended by Weeder adviser. All promoters that were available through Phytozome BioMart from all species included in the analyses, was used as a background set (a total of 1029220 promoters). A separate parser was written to extract significant PFMs from result files produced by each program. The PFMs produced for each dehydrin class were entered into the STAMP [85] website to group matrices by similarity and to identify significant (E-value ≤ 0.05) matches in PLACE [23]. A representative member from a tree node of matrices grouped by similarity was selected and its sequence logo was generated using WebLogo 3.3 [86].

Results and Discussion

In order to further understand how different dehydrins are regulated in response to environmental stress, motifs corresponding to conserved *cis*-regulatory elements were detected in the upstream regions of dehydrin genes in all five subclasses. Dehydrin proteins are by nature unstructured, and a custom identification strategy was employed to retrieve as many dehydrin genes with up to 1000 bp upstream region as possible. In total, 340 dehydrin promoters of size 1000 bp and eight dehydrin promoters of shorter length were retrieved from 51 plant genome sequences. In addition, two promoters from dehydrin genes isolated from two *Oxytropis* species were also included (Table 1, S1 Table). Out of the queried genomes, 10 were from monocotyledonous plants, 37 from dicotyledonous plants, one from a basal angiosperm (*Amborella trichopoda*), two from gymnosperms (*Picea abies* and *Pinus taeda*) and one from moss (*Physcomitrella patens*). The 350 sequences identified were confirmed to also match annotated dehydrins. For 330 out of 350 sequences, at least half of the top ten significant hits had "Dehydrin" as domain annotation. For the remaining 20 putative dehydrins, less than half of the top ten significant hits carried that annotation. Out of those, three were annotated as either a dehydrin or similar to dehydrin on NCBI GenBank and two were annotated as having a dehydrin

Table 1. Number of analyzed dehydrin promoters per species.

Species	K_n	KS	SK_n	Y_nK_n	Y_nSK_n	Total
Actinidia chinensis	2	1	4	0	2	9
Amborella trichopoda	0	1	2	0	0	3
Aquilegia coerulea	1	1	2	1	1	6
Arabidopsis halleri	0	0	3	1	2	6
Arabidopsis lyrata	1	1	5	1	2	10
Arabidopsis thaliana	1	1	4	1	3	10
Beta vulgaris	0	0	2	0	2	4
Boechera stricta	1	1	4	0	3	9
Brachypodium distachyon	0	1	4	0	5	10
Brassica rapa	0	1	6	0	4	11
Cajanus cajan	0	1	2	1	1	5
Capsella grandiflora	1	1	6	0	1	9
Capsella rubella	1	1	4	0	3	9
Capsicum annuum	0	2	2	0	5	9
Carica papaya	0	1	1	0	2	4
Cicer arietinum	0	2	1	1	1	5
Citrus clementina	1	0	1	0	2	4
Citrus sinensis	0	1	2	0	3	6
Cucumis sativus	0	1	1	0	2	4
Eucalyptus grandis	1	0	1	0	4	6
Eutrema salsugineum	0	1	4	1	3	9
Fragaria vesca	0	0	1	0	5	6
Glycine max	0	4	1	2	3	10
Gossypium raimondii	1	0	3	1	3	8
Linum usitatissimum	0	1	6	2	2	11
Lotus japonicus	0	0	2	1	1	4
Malus domestica	0	0	2	2	5	9
Manihot esculenta	0	2	1	0	2	5
Medicago truncatula	0	1	1	1	1	4
Mimulus guttatus	0	1	1	0	2	4
Musa acuminata	0	2	1	0	1	4
Oryza brachyantha	0	1	1	0	5	7
Oryza sativa	0	1	1	0	6	8
Panicum virgatum	0	3	2	0	5	10
Phaseolus vulgaris	0	1	1	1	1	4
Phoenix dactylifera	2	1	2	0	1	6
Physcomitrella patens	4	0	0	0	0	4
Picea abies	4	0	9	0	0	13
Pinus taeda	8	0	10	0	0	18
Populus trichocarpa	3	1	1	0	2	7
Prunus persica	0	0	2	2	2	6
Ricinus communis	0	1	1	0	3	5
Salix purpurea	5	1	2	2	0	10
Setaria italica	0	0	2	0	5	7
Solanum lycopersicum	0	0	1	0	4	5
Solanum tuberosum	0	1	1	0	3	5
Sorghum bicolor	0	1	1	0	3	5

(Continued)

Table 1. (*Continued*)

Species	K_n	KS	SK_n	Y_nK_n	Y_nSK_n	Total
Spirodela polyrhiza	1	0	0	0	0	1
Theobroma cacao	1	1	1	0	2	5
Vitis vinifera	0	0	0	0	2	2
Zea mays	0	2	2	0	3	7
*Oxytropis splendens**	–	1	–	–	–	1
*Oxytropis arctobia**	–	1	–	–	–	1
Total	39	47	120	21	123	350

* Only one promoter was obtained per species, using genome walking.

doi:10.1371/journal.pone.0129016.t001

domain on UniProtKB. The rest of the sequences were all short putative KS dehydrins. In these cases, all significant phmmer hits were analyzed. From 14.2% to 30.1% of significant hits had "Dehydrin" domain architecture, for sequences with lowest and highest number of significant hits with "Dehydrin" annotation, respectively. The rest of significant hits had no architecture annotation. When Needleman-Wunsch pairwise alignment was used to compare 15 putative KS dehydrin to a known *Arabidopsis* KS dehydrin (AT1G54410), sequence similarities ranged from 59.0% to 98.0%. This evidence supports the notion that the sequences extracted for the analyses can be classified as dehydrins.

Biochemical properties of dehydrins

Dehydrins are known to be intrinsically disordered and hydrophilic [87], making it difficult, if not impossible, to identify them by overall sequence homology. These properties are, however, important for their hypothesized function in protein stabilization through interaction with water molecules, as well as for their subcellular location in the cytosol and the nucleus and not within membranes [87,88]. To assess the identification of the dehydrins included in this study, the grand average of hydropathicity (GRAVY, [89]) and the proportions of amino acids in the disordered regions were compared between dehydrins and random plant proteins. It was found that the 350 dehydrin amino acid sequences analyzed, were significantly more hydrophilic than 350 random plant protein sequences (GRAVY -1.3470 for dehydrins, -0.2938 for random plant proteins, p-value < 0.001). The level of structural disorder indicated that in the dehydrins analyzed, the average proportion of amino acid sequences in the state of disorder was 99.32% compared to 15.95% in random plant proteins (p-value < 0.001).

Promoters of KS dehydrins have one conserved GATA motif

In total, 47 KS dehydrin promoters were included in the *de novo* motif discovery (Table 1). Using the *de novo* motif discovery tool Seeder [83,90], one single putative conserved regulatory motif was discovered in all 47 promoter sequences (Motif 1, Table 2, S2 Table). A similar motif was also discovered with Weeder [84]. The KS dehydrins are known to be expressed in response to cold and dehydration, as well as being constitutively expressed [6,91–93]. Although the single identified overrepresented motif in KS dehydrin promoters does not directly match any typical cold or dehydration-related *cis*-regulatory elements in the PLACE database [23], it does match two motifs involved in light regulation and one involved in sugar regulation (Table 2): IBOXCORENT (I-box core) [94], REBETALGLHCB21 [95] and SREATMSD [96], respectively. These three experimentally validated motifs share four nucleotides (GATA).

Table 2. Selected *de novo* motifs found in KS dehydrin promoters and their putative function identified through PLACE database.

		Match in PLACE[5]		
De novo motif	Sequence	Name	E-value[6]	Function
1. Seeder[1] AWTCGGATAA[2] (47/47[3], 8.8e-07[4])	GATAAGR	IBOXCORENT	1.4e-08	Found in light-responsive conserved DNA modular arrays
	TTATCC	SREATMSD	1.2e-07	Sugar-repressive element (SRE) found in genes down-regulated after main stem decapitation
	CGGATA	REBETALGLHCB21	1.8e-07	Required for phytochrome regulation

[1]Number of the motif and the *de novo* discovery software that was used to locate that motif.

[2]Motif consensus sequence in IUPAC nucleotide code.

[3]Occurrence is the number of promoters containing a *de novo* motif out of the total number of promoters analyzed for a specific dehydrin class, presented in the parentheses.

[4]Siginificance of the motif, E-value calculated by MEME, Q-value calculated by Seeder, presented in the parentheses.

[5]PLACE matches were identified using STAMP, only significant matches with E-value < 0.05 are presented.

[6]E-value of the match with PLACE motif.

doi:10.1371/journal.pone.0129016.t002

One of these motifs, the I-box (GATAAGR) can form a light-responsive conserved DNA modular array (CMA) together with a G-box (CACGTGGC) when located in close proximity to one another. In transgenic Arabidopsis and tobacco (*Nicotiana tabacum*) plants, the presence of this CMA in a promoter, drives GUS reporter gene expression when exposed to light. Interestingly, this expression seems to be mediated by phytochrome and cryptochrome photoreceptors [94].

Another of the motifs matching the motif discovered in the KS dehydrin promoters, the REBETALGLHCB21, also called REβ (CGGATA), was first identified in gibbous duckweed (*Lemna gibba*) [95]. It is involved in phytochrome-mediated repression of promoter activity in darkness, when located in close proximity with REα (AACCAA). Although REα was not identified as a significantly overrepresented motif, it is found in 26 out of the 47 KS dehydrin promoters analyzed. The GATA part of the REβ was shown to be absolutely necessary for darkness-induced repression [95]. Furthermore, in Arabidopsis, C-repeat (CCGAC, CRT)-linked cold and dehydration induced gene expression is mediated by phytochrome [25]. While CRT was not found to be significantly overrepresented within the set of KS dehydrin promoters, it is noteworthy that 27 out of the 47 KS dehydrin promoters contain one or more copies of CRT or its reverse complement. Sixteen of the promoters contain both REα and REβ.

The motif discovered in the KS dehydrin promoters also matched a sugar-repressive element, SREATMSD (TTATCC, SRE), shown to be involved in sugar mediated gene repression in Arabidopsis [96]. Sugars are known osmoprotectants that are produced by plants in response to cold [97]. One of the suggested roles of dehydrins is in the stabilization of protein conformation. Sugars, such as sucrose and trehalose, can replace water molecules on the surface of a protein and can thus conserve its conformation. This allows cells to restore their function after rehydration [98].

Motifs discovered in promoters of K_n match abscisic acid and low temperature response elements

A total of 39 K_n dehydrin promoters were included in the *de novo* regulatory motif discovery analysis (Table 1). The K_n dehydrins are expressed in response to high salinity, abscisic acid (ABA), cold and dehydration [3,5,99,100]. A total of three putative regulatory motifs were identified in this set of promoters (Table 3, S2 Table)—two were discovered using MEME

Table 3. Selected *de novo* motifs found in K_n dehydrin promoters and their putative function identified through PLACE database.

| *De novo* motif | Match in PLACE[5] | | | |
	Sequence	Name	E-value[6]	Function
2. MEME[1] GGCMCCAC[2] (31/39[3], 1.2e-06[4])	TGGTCCCAC	SITEIIBOSPCNA	4.3e-07	Involved for meristematic tissue-specific expression in rice
3. MEME AYGTCGGY (39/39, 3.7e-05)	TACCGACAT	DREDR1ATRD29AB	6.4e-11	Response to drought, low temperature and high salinity. Bound by CBF1 in *Arabidopsis*
	ACCGACA	LTREATLTI78	5.5e-09	LTRE
	ACGTSSSC	ABREOSRAB21	3.3e-05	ABRE found in wheat and rice
4. Seeder WNRCCGACAT (39/39, 2.1e-05)	ACCGACA	LTREATLTI78	3.7e-07	LTRE
	TACCGACAT	DREDR1ATRD29AB	4.9e-08	Response to drought, low temperature and high salinity. Bound by CBF1 in *Arabidopsis*

[1]Number of the motif and the *de novo* discovery software that was used to locate that motif.

[2]Motif consensus sequence in IUPAC nucleotide code.

[3]Occurrence is the number of promoters containing a *de novo* motif out of the total number of promoters analyzed for a specific dehydrin class, presented in the parentheses.

[4]Siginificance of the motif, E-value calculated by MEME, Q-value calculated by Seeder, presented in the parentheses.

[5]PLACE matches were identified using STAMP, only significant matches with E-value < 0.05 are presented.

[6]E-value of the match with PLACE motif.

doi:10.1371/journal.pone.0129016.t003

(Motif 2: GGCAGGAC/GTGGTGCC; and Motif 3: ATGTCGGC/GCCGACAT) and one using Seeder (Motif 4: TCGCCGACAT/ATGTCGGCGA). Motif 2 (GGCAGGAC) has a significant match to the SITEIIBOSPCNA (TGGTCCCAC) motif in the PLACE database. This motif is linked with meristematic tissue-specific gene expression in rice (*Oryza sativa*) [101] and it was found in 31 out of the 39 promoters. Motifs 3 and 4, found in all analyzed K_n dehydrin promoters, match DREDR1ATRD29AB motif (TACCGACAT) [102] and LTREATLTI78 (ACCGACA) [103], two low temperature response elements (LTREs) involved in cold response in *A. thaliana*. Additionally, Motif 3 matches an ABRE found in wheat and rice- ABREOSRAB21 (ACGTSSSC) [104]. The presence of both LTREs and an ABRE indicates that K_n dehydrins, similarly to SK_n and Y_nSK_n dehydrins, could be expressed in ABA-dependent and independent manner in response to osmotic stresses.

SK_n dehydrins contain multiple cold/dehydration, abscisic acid and light regulated response elements

A total of 120 SK_n dehydrin promoters were analyzed (Table 1). Six *de novo* discovered putative regulatory motifs are presented in Table 4 and S2 Table. MEME and Seeder each discovered three motifs. The SK_n dehydrins are known to be expressed in response to cold, ABA, dehydration and salt [3,5,14,105]. Three out of six motifs (motifs 5–7) have matches in PLACE that are known ABREs. Motif 5 (CCACGTGTC/GACACGTGG) matches ABREs from wheat (*Triticum aestivum*) [103] and canola (*Brassica napus*) [106]. Motif 6 (CCGACGCG/ CGCGTCGG) matches ABREs from maize [107], and rice [108]. Motif 7 (CCAACGCG/ CGCGTTGG) matches an ABRE from barley [109] and rice [107]. Motifs 6, 8 (CACCGACC/ GGTCGGTG) and 9 (TGGTCGGT/ACCGACCA) match low temperature response elements known as C-repeats (CRT, consensus sequence: RCCGAC), found in numerous species [110–112].

Table 4. Selected *de novo* motifs found in SK_n dehydrin promoters and their putative function identified through PLACE database.

De novo motif	Match in PLACE[5]			
	Sequence	Name	E-value[6]	Function
5. Seeder[1] MCACGTGTC[2] (120/120[3], 9.2e-19[4])	GGACACGTGGC	ABRETAEM	6.5e-13	ABRE found in wheat
	TGACACGTGGCA	HY5AT	2.4e-12	Bound by HY5, involved in light regulation of transcriptional activity
	TCCACGTGTC	SGBFGMGMAUX28	1.1e-13	Recognized by G-box binding factors in soybean. Found in auxin-responsive genes
	MCACGTGGC	GBOXLERBCS	8.1e-12	Sequence found in promoters of light-regulated genes
	CGCCACGTGTCC	ABREBNNAPA	2.4e-21	ABRE found in *Brassica napus*
6. MEME CCGACGCG (120/120, 1.2e-50)	CCCACGTGGC	ABREAZMRAB28	3.0e-05	ABRE, ABA and water-stress responses. Binding site of CBF2.
	GCCGCGTGGC	ABREMOTIFIIIOSRAB16B	4.0e-05	ABRE Motif III found in rice
	CCACGTGGCC	LTRECOREATCOR15	5.3e-06	LTRE
7. Seeder SCAACGCG (120/120, 2.0e-10)	TCCACGTCTC	ABRE3HVA1	2.3e-05	ABRE found in barley
	GCCGCGTGGC	ABREMOTIFIIIOSRAB16B	6.0e-05	ABRE Motif III found in rice
8. MEME CACCGACC (119/120, 3.8e-59)	RCCGAC	DRECRTCOREAT	1.9e-08	DRE/CRT found in genes expressed in response to cold and dehydration
	CCCACCTACC	ACIPVPAL2	3.8e-07	Required for vascular specific expression
	ACCGACA	LTREATLTI78	3.1e-07	LTRE
9. Seeder KKGTCGGY (120/120, 4.9e-07)	ACCGACA	LTREATLTI78	8.4e-07	LTRE
	RCCGAC	DRECRTCOREAT	2.4e-08	DRE/CRT found in genes expressed in response to cold and dehydration
10. MEME GTGGGVCC (61/120, 2.5e-22)	GGTCCCAT	GGTCCCATGMSAUR	8.3e-09	Auxin RE found in soybean
	CTCCCAC	BOXCPSAS1	3.9e-06	Involved in light-induced repression

[1]Number of the motif and the *de novo* discovery software that was used to locate that motif.

[2]Motif consensus sequence in IUPAC nucleotide code.

[3]Occurrence is the number of promoters containing a *de novo* motif out of the total number of promoters analyzed for a specific dehydrin class, presented in the parentheses.

[4]Siginificance of the motif, E-value calculated by MEME, Q-value calculated by Seeder, presented in the parentheses.

[5]PLACE matches were identified using STAMP, only significant matches with E-value < 0.05 are presented.

[6]E-value of the match with PLACE motif.

doi:10.1371/journal.pone.0129016.t004

In addition, motif 5 matches an element from tomato and Arabidopsis light-regulated genes [113,114] and motif 10 (GTGGGACC) matches an element from pea involved in light-induced repression [115].

The presence of these significantly overrepresented motifs indicates that the SK_n dehydrins are regulated at the transcriptional level and their expression is modulated in response to cold and ABA. SK_n dehydrins should also be expressed in response to drought, since CRT, which is also called dehydration responsive element [116], is found in their promoters. The circadian clock controls cold induction of C-repeat binding factors (CBFs), which in turn bind CRT/DRE elements [117]. Phytochrome and cryptochrome genes are also regulated by a circadian clock in Arabidopsis [118]. COR27, a cold-induced gene, is regulated by circadian clock related evening elements (EE) [119]. In addition to EE, the COR27 promoter also contains multiple ABREs and G-boxes, to which motifs 5 and 6 also match. The core EE (AATATCT) [120] is

found in 18 out of 73 SK_n gene promoters analyzed. Motifs involved in light-induced regulation of gene expression found in the promoters of SK_n genes could participate in modulation of these genes by the circadian clock. It has been shown previously, using bioinformatics methods, that the promoters of cold-regulated genes contain CRTs and ABREs [112,121] and our data also support those findings.

Motifs 5 and 10 match an auxin response element found in soybean GmAux28 [122] and SAUR15A promoter, respectively [123]. It has been shown previously that numerous genes related to auxin response in Arabidopsis are modulated in response to cold, such as auxin response factor ARF7 or the PINOID-binding protein 1 that is involved in hormone signaling and stress response [124].

Y_nSK_n dehydrins promoters contain multiple ABREs, light REs and a CRT

Y_nSK_n dehydrins represent the largest subclass out of the five dehydrin classes analyzed. A total of 123 Y_nSK_n gene promoters were analyzed ([Table 1]). The Y_nSK_n dehydrins are expressed in response to ABA, dehydration and high salinity [3,5,100]. The two motifs presented in [Table 5] and [S2 Table] (Motifs 11 and 12) match numerous elements in the PLACE database and both were discovered using a Seeder. Motif 11 (GACACGTGGC) is very similar to Motif 5 (GACACGTGT), found in the SK_n dehydrin promoters and they both match several of the same motifs in PLACE database, namely ABREs, G-box and light response elements. Motif 13 (CACCGAC) is almost identical to Motif 8 (CACCGACC) discovered in the SK_n dehydrin promoters, which matches CRT/DRE necessary for CBF mediated cold and dehydration response [116]. Overall, motifs found Y_nSK_n dehydrin promoters are very similar to those found in SK_n dehydrin promoters indicating that they possibly have a similar function, and that these two classes may have diverged more recently than the other classes. While the function of the Y-segment in the gene products of Y_nSK_n and Y_nK_n dehydrins is not known, it shows similarity

Table 5. Selected *de novo* motifs found in Y_nSK_n dehydrin promoters and their putative function identified through PLACE database.

De novo motif	Match in PLACE[5]			
	Sequence	Name	E-value[6]	Function
11. Seeder[1] SACACGTGG[2] (123/123[3], 4.8e-38[4])	GGACACGTGGC	ABRETAEM	1.1e-16	ABRE found in wheat
	TCCACGTGTC	SGBFGMGMGAUX28	2.5e-13	Recognized by G-box binding factors in soybean. Found in auxin-responsive genes
	CGCCACGTGTCC	ABREBNNAPA	6.7e-16	ABRE found in rapeseed
	TGACACGTGGCA	HY5AT	6.7e-16	Bound by HY5, involved in light regulation of transcriptional activity
	MCACGTGGC	GBOXLERBCS	7.3e-14	Sequence found in promoters of light-regulated genes
12. Seeder CRCCGAC (123/123, 3.1e-11)	RCCGAC	DRECRTCOREAT	2.3e-09	DRE/CRT found in gene expressed in response to cold and dehydration
	RYCGAC	CBFHV	1.4e-07	CRT found in barley (*Hordeum vulgare*)

[1]Number of the motif and the *de novo* discovery software that was used to locate that motif.

[2]Motif consensus sequence in IUPAC nucleotide code.

[3]Occurrence is the number of promoters containing a *de novo* motif out of the total number of promoters analyzed for a specific dehydrin class, presented in the parentheses.

[4]Siginificance of the motif, E-value calculated by MEME, Q-value calculated by Seeder, presented in the parentheses.

[5]PLACE matches were identified using STAMP, only significant matches with E-value < 0.05 are presented.

[6]E-value of the match with PLACE motif.

doi:10.1371/journal.pone.0129016.t005

to the nucleotide binding domain of plant chaperones [125]. The gene products of the other dehydrin classes do not have any such domains. In addition, there are evolutive constraints on the Y-segment in a dehydrin from arctic *Oxytropis* species compared with temperate species [126], suggesting that the Y-segment might carry an important function that differentiates Y_nSK_n from SK_n dehydrins. Some of the published data shows that Y_nSK_n dehydrins are not expressed in response to cold [3,5], however there is evidence that after a period of acclimation they do accumulate in Red-Osier Dogwood (*Cornus sericea* L.) [127] and apple trees [128]. It is possible that cold-induced Y_nSK_n dehydrin expression was not detected in some data sets due to a limited time of exposure to low temperature.

Y_nK_n dehydrins promoters contain ABREs and light REs

Y_nK_n dehydrins represent the smallest subgroup, with only 21 members found in 51 plant genomes (Table 1). Y_nK_n dehydrins are known to be expressed in response to cold [129,130], and two motifs were detected in their promoters (Table 6, S2 Table). One was identified using Seeder and the other using MEME. Both motifs match several ABREs and light REs in the PLACE database. Motif 13 (TAACACGTGTC/GACACGTGTTA) is similar to motif 11 (GACACGTGGC/GCCACGTGTC) identified in Y_nSK_n dehydrins and it matches the some of the same motifs in PLACE. Motif 14 (ACGTGGCA/TGCCACGT) is similar to motif 11 (GACACGTGGC) found in Y_nSK_n dehydrin. The lack of CRTs in the promoters of Y_nK_n dehydrins suggests that they might be expressed in response to cold in ABA-dependent manner, not linked with the CBF transcription factors [18].

Table 6. Selected *de novo* motifs found in Y_nK_n dehydrin promoters and their putative function identified through PLACE database.

De novo motif	Sequence	Name	E-value[6]	Function
		Match in PLACE[5]		
13. Seeder[1] YRACACGTGTCC[2] (21/21[3], 1.5e-09[4])	GGACACGTGGC	ABRETAEM	4.1-e11	ABRE found in wheat
	CGCCACGTGTCC	ABREBNNAPA	7.8e-11	ABRE found in rapeseed
	CGCACGTGTC	ABRE2HVA22	6.3e-09	ABRE2 found in barley HVA22 gene
	TCCACGTGTC	SGBFGMGMAUX28	6.3e-09	Recognized by G-box binding factors in soybean. Found in auxin-responsive genes
	AACGCGTGTC	CE3OSOSEM	1.1e-10	Coupling element 3 found in rice, required for ABA induced expression
	TGACACGTGGCA	HY5AT	7.5e-09	Bound by HY5, involved in light regulation of transcriptional activity
14. MEME ACGTGKCA (21/21, 8.3e-05)	ACGTGKC	ACGTABREMOTIFA2OSEM	1.3e-11	Core of ABRE in rice
	ACGTGGCA	LRENPCABE	5.7e-13	Positive light RE in tobacco
	YACGTGGC	ABREATCONSENSUS	7.4e-09	ABRE found in Arabidopsis
	RTACGTGGCR	ABADESI1	2.0e-11	ABRE and desiccation response in rice

[1]Number of the motif and the *de novo* discovery software that was used to locate that motif.

[2]Motif consensus sequence in IUPAC nucleotide code.

[3]Occurrence is the number of promoters containing a *de novo* motif out of the total number of promoters analyzed for a specific dehydrin class, presented in the parentheses.

[4]Siginificance of the motif, E-value calculated by MEME, Q-value calculated by Seeder, presented in the parentheses.

[5]PLACE matches were identified using STAMP, only significant matches with E-value < 0.05 are presented.

[6]E-value of the match with PLACE motif.

doi:10.1371/journal.pone.0129016.t006

Conclusions

Numerous dehydrins were identified in 51 plant genomes, many of which are not found in protein databases such as InterPro or PROSITE, or they are not annotated in Phytozome. The identified dehydrins were categorized into five subclasses based on the occurrence of conserved protein segments. Three *de novo* motif discovery software tools were used to find statistically significant overrepresented motifs in the promoters of each group of dehydrins. These motifs were matched to known *cis*-regulatory elements in the PLACE database to help explain the regulation of dehydrin expression in response to different environmental stimuli.

Dehydrins are expressed in response to multiple stress stimuli. Although there is overlap in expression triggers between dehydrin subclasses, there are differences in the pattern of expression. Some of the dehydrins are expressed constitutively in all tissues [3,5] and more specifically in seeds [131,132]. The presence of ABREs, CRTs and light REs in the promoters of Y_nSK_n and SK_n dehydrins indicates that they could be expressed in response to dehydration and cold in both ABA-dependent and independent pathway and that this expression is modulated by light.

While Y_nSK_n and SK_n dehydrin are found in most species, often in several copies, the other three subclasses are encountered less often. It is probable that they either have specialized functions or they are expressed together with Y_nSK_n and SK_n dehydrins to increase the overall protective effect against dehydration. It is important to note that the number of discovered dehydrins is probably an underestimation due to incompleteness of genome assembly and errors inherent in sequencing.

Dehydrins play an important role in the survival of plants facing various stresses. Motifs matching *cis*-regulatory elements linked to both ABA-dependent and independent stress response pathways, as well as light response pathways were detected in dehydrins from many different plant families. The implication of this finding is that the regulation of dehydrins is conserved in the plant lineages included in this study and that stress-linked selection pressure preserved *cis*-regulatory elements in the promoters of dehydrins through stabilizing selection.

Supporting Information

S1 Table. Annotation and meta-data about the dehydrins included in the study. Each identified dehydrin was further analyzed by BLAST to find the closest match at NCBI GenBank non-reduntant database. The fields are 1. Species; 2. Gene; 3. Dehydrin subgroup; 4. BLAST top hit e-value; 5. BLAST top hit accession; 6. BLAST top hit description; 7. K-segment location; 8. Y-segment location; 9. S-segment location.
(CSV)

S2 Table. Motif logos of motifs discovered in dehydrin promoters. Weblogos were made for each of the motifs identified by *de novo* motif discovery algorithms in five classes of dehydrin genes. The motif numbers correspond to the motif numbers in Tables 2–6.
(DOCX)

Acknowledgments

The authors wish to thank Paul Harrison for helpful discussions on disordered proteins.

Author Contributions

Conceived and designed the experiments: MS YZ. Performed the experiments: YZ. Analyzed the data: YZ. Contributed reagents/materials/analysis tools: MS. Wrote the paper: MS YZ.

References

1. Hannah MA, Heyer AG, Hincha DK (2005) A global survey of gene regulation during cold acclimation in *Arabidopsis thaliana*. PLoS Genet 1: e26. doi: 10.1371/journal.pgen.0010026 PMID: 16121258

2. Maruyama K, Takeda M, Kidokoro S, Yamada K, Sakuma Y, Urano K, et al. (2009) Metabolic Pathways Involved in Cold Acclimation Identified by Integrated Analysis of Metabolites and Transcripts Regulated by DREB1A and DREB2A. Plant Physiol 150: 1972–1980. doi: 10.1104/pp.109.135327 PMID: 19502356

3. Choi DW, Close TJ, Zhu B (1999) The barley (*Hordeum vulgare* L.) dehydrin multigene family: sequences, allele types, chromosome assignments, and expression characteristics of 11 Dhn genes of cv Dicktoo. Theor Appl Genet 98: 1234–1247–1247. doi: 10.1007/s001220051189

4. Allagulova CR, Gimalov FR, Shakirova FM, Vakhitov VA (2003) The Plant Dehydrins: Structure and Putative Functions. Biochemistry (Moscow) 68: 945–951. doi: 10.1023/A:1026077825584 PMID: 14606934

5. Hundertmark M, Hincha DK (2008) LEA (late embryogenesis abundant) proteins and their encoding genes in *Arabidopsis thaliana*. BMC Genomics 9: 118. doi: 10.1186/1471-2164-9-118 PMID: 18318901

6. Archambault A, Strömvik MV (2011) PR-10, defensin and cold dehydrin genes are among those over expressed in *Oxytropis* (Fabaceae) species adapted to the arctic. Funct Integr Genomic 11: 497–505. doi: 10.1007/s10142-011-0223-6 PMID: 21499864

7. Yamasaki Y, Koehler G, Blacklock BJ, Randall SK (2013) Dehydrin expression in soybean. Plant Physiol and Bioch 70: 213–220. doi: 10.1016/j.plaphy.2013.05.013 PMID: 23792826

8. Close TJ (1996) Dehydrins: Emergence of a biochemical role of a family of plant dehydration proteins. Physiol Plant 97: 795–803. doi: 10.1111/j.1399-3054.1996.tb00546.x

9. Rorat T (2006) Plant dehydrins—tissue location, structure and function. Cell Mol Biol Lett 11: 536–556. doi: 10.2478/s11658-006-0044-0 PMID: 16983453

10. Koag MC, Wilkens S, Fenton RD, Resnik J, Vo E, Close TJ (2009) The K-Segment of Maize DHN1 Mediates Binding to Anionic Phospholipid Vesicles and Concomitant Structural Changes. Plant Physiol 150: 1503–1514. doi: 10.1104/pp.109.136697 PMID: 19439573

11. Drira M, Saibi W, Brini F, Gargouri A, Masmoudi K, Hanin M (2012) The K-Segments of the Wheat Dehydrin DHN-5 are Essential for the Protection of Lactate Dehydrogenase and β-Glucosidase Activities *In Vitro*. Mol Biotechnol 54: 643–650. doi: 10.1007/s12033-012-9606-8

12. Rahman LN, McKay F, Giuliani M, Quirk A, Moffatt BA, Harauz G, et al. (2013) Interactions of *Thellungiella salsuginea* dehydrins TsDHN-1 and TsDHN-2 with membranes at cold and ambient temperatures—Surface morphology and single-molecule force measurements show phase separation, and reveal tertiary and quaternary associations. BBA—Biomembranes 1828: 967–980. doi: 10.1016/j.bbamem.2012.11.031

13. Close TJ (1997) Dehydrins: A commonalty in the response of plants to dehydration and low temperature. Physiol Plant 100: 291–296. doi: 10.1111/j.1399-3054.1997.tb04785.x

14. Kovacs D, Kalmar E, Torok Z, Tompa P (2008) Chaperone Activity of ERD10 and ERD14, Two Disordered Stress-Related Plant Proteins. Plant Physiol 147: 381–390. doi: 10.1104/pp.108.118208 PMID: 18359842

15. Rahman LN, Smith GST, Bamm VV, Voyer-Grant JAM, Moffatt BA, Dutcher JR, et al. (2011) Phosphorylation of *Thellungiella salsuginea* dehydrins TsDHN-1 and TsDHN-2 facilitates cation-induced conformational changes and actin assembly. Biochemistry 50: 9587–9604. doi: 10.1021/bi201205m PMID: 21970344

16. Eriksson SK, Harryson P (2011) Dehydrins: Molecular Biology, Structure and Function. In: Lüttge U, Beck E, Bartels D, editors. Ecological Studies. Ecological Studies. Berlin, Heidelberg: Springer Berlin Heidelberg, Vol. 215. pp. 289–305. doi: 10.1007/978-3-642-19106-0_14

17. Abba S, Ghignone S, Bonfante P (2006) A dehydration-inducible gene in the truffle *Tuber borchii* identifies a novel group of dehydrins. BMC Genomics 7: 39. doi: 10.1186/1471-2164-7-39 PMID: 16512918

18. Agarwal PK, Jha B (2010) Transcription factors in plants and ABA dependent and independent abiotic stress signalling. Biol Plant 54: 201–212. doi: 10.1007/s10535-010-0038-7

19. Sakuma Y, Liu Q, Dubouzet JG, Abe H, Shinozaki K, Yamaguchi-Shinozaki K (2002) DNA-Binding Specificity of the ERF/AP2 Domain of Arabidopsis DREBs, Transcription Factors Involved in Dehydration- and Cold-Inducible Gene Expression. Biochem Bioph Res Co 290: 998–1009. doi: 10.1006/bbrc.2001.6299

20. Knight H, Trewavas AJ, Knight MR (1997) Calcium signalling in *Arabidopsis thaliana* responding to drought and salinity. Plant J 12: 1067–1078. doi: 10.1046/j.1365-313X.1997.12051067.x PMID: 9418048

21. Thomashow MF, Gilmour SJ, Stockinger EJ, Jaglo-Ottosen KR, Zarka DG (2001) Role of the Arabidopsis CBF transcriptional activators in cold acclimation. Physiol Plant 112: 171–175. doi: 10.1034/j.1399-3054.2001.1120204.x

22. Fowler S, Thomashow MF (2002) Arabidopsis Transcriptome Profiling Indicates That Multiple Regulatory Pathways Are Activated during Cold Acclimation in Addition to the CBF Cold Response Pathway. Plant Cell 14: 1675–1690. doi: 10.1105/tpc.003483 PMID: 12172015

23. Higo K (1999) Plant cis-acting regulatory DNA elements (PLACE) database: 1999. Nucleic Acids Res 27: 297–300. doi: 10.1093/nar/27.1.297 PMID: 9847208

24. Chinnusamy V, Schumaker K, Zhu J-K (2004) Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. J Exp Bot 55: 225–236. doi: 10.1093/jxb/erh005 PMID: 14673035

25. Kim H-J, Kim Y-K, Park J-Y, Kim J (2002) Light signalling mediated by phytochrome plays an important role in cold-induced gene expression through the C-repeat/dehydration responsive element (C/DRE) in *Arabidopsis thaliana*. Plant J 29: 693–704. doi: 10.1046/j.1365-313X.2002.01249.x PMID: 12148528

26. Lee C-M, Thomashow MF (2012) Photoperiodic regulation of the C-repeat binding factor (CBF) cold acclimation pathway and freezing tolerance in *Arabidopsis thaliana*. P Natl Acad Sci USA 109: 15054–15059. doi: 10.1073/pnas.1211295109 PMID: 22927419

27. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40: D1178–D1186. doi: 10.1093/nar/gkr944 PMID: 22110026

28. Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, et al. (2009) BioMart—biological queries made easy. BMC Genomics 10: 22. doi: 10.1186/1471-2164-10-22 PMID: 19144180

29. Amborella Genome Project (2013) The Amborella genome and the evolution of flowering plants. Science 342: 1241089. doi: 10.1126/science.1241089 PMID: 24357323

30. Hu TT, Pattyn P, Bakker EG, Cao J, Cheng J-F, Clark RM, et al. (2011) The Arabidopsis lyrata genome sequence and the basis of rapid genome size change. Nat Genet 43: 476–481. doi: 10.1038/ng.807 PMID: 21478890

31. The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408: 796–815. doi: 10.1038/35048692 PMID: 11130711

32. The International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463: 763–768. doi: 10.1038/nature08747 PMID: 20148030

33. Cheng F, Liu S, Wu J, Fang L, Sun S, Liu B, et al. (2011) BRAD, the genetics and genomics database for *Brassica plants*. BMC Plant Biol 11: 136. doi: 10.1186/1471-2164-12-239 PMID: 21995777

34. The *Brassica rapa* Genome Sequencing Project Consortium (2011) The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet 43: 1035–1039. doi: 10.1038/ng.919 PMID: 21873998

35. Ming R, Hou S, Feng Y, Yu Q, Dionne-Laporte A, Saw JH, et al. (2008) The draft genome of the transgenic tropical fruit tree papaya (*Carica papaya* Linnaeus). Nature 452: 991–996. doi: 10.1038/nature06856 PMID: 18432245

36. Slotte T, Hazzouri KM, Ågren JA, Koenig D, Maumus F, Guo Y-L, et al. (2013) The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. Nat Genet 45: 831–835. doi: 10.1038/ng.2669 PMID: 23749190

37. Wu GA, Prochnik S, Jenkins J, Salse J, Hellsten U, Murat F, et al. (2014) Sequencing of diverse mandarin, pummelo and orange genomes reveals complex history of admixture during citrus domestication. Nat Biotechnol 32: 656–662. doi: 10.1038/nbt.2906 PMID: 24908277

38. Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, Grimwood J, et al. (2014) The genome of *Eucalyptus grandis*. Nature 510: 356–362. doi: 10.1038/nature13308 PMID: 24919147

39. Yang R, Jarvis DE, Chen H, Beilstein MA, Grimwood J, Jenkins J, et al. (2013) The Reference Genome of the Halophytic Plant *Eutrema salsugineum*. Front Plant Sci 4. doi: 10.3389/fpls.2013.00046

40. Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, et al. (2010) The genome of woodland strawberry (*Fragaria vesca*). Nat Genet 43: 109–116. doi: 10.1038/ng.740 PMID: 21186353

41. Schmutz J, Cannon SB, Schlueter J, Ma J, Mitros T, Nelson W, et al. (2010) Genome sequence of the palaeopolyploid soybean. Nature 463: 178–183. doi: 10.1038/nature08670 PMID: 20075913

42. Paterson AH, Wendel JF, Gundlach H, Guo H, Jenkins J, Jin D, et al. (2012) Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature 492: 423–427. doi: 10.1038/nature11798 PMID: 23257886

43. Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, et al. (2012) The genome of flax (*Linum usitatissimum*) assembled de novo from short shotgun sequence reads. Plant J 72: 461–473. doi: 10.1111/j.1365-313X.2012.05093.x PMID: 22757964

44. Velasco R, Zharkikh A, Affourtit J, Dhingra A, Cestaro A, Kalyanaraman A, et al. (2010) The genome of the domesticated apple (*Malus* × *domestica* Borkh.). Nat Genet 42: 833–839. doi: 10.1038/ng.654 PMID: 20802477

45. Prochnik S, Marri PR, Desany B, Rabinowicz PD, Kodira C, Mohiuddin M, et al. (2012) The Cassava Genome: Current Progress, Future Directions. Trop Plant Biol 5: 88–94. doi: 10.1007/s12042-011-9088-z PMID: 22523606

46. Young ND, Debellé F, Oldroyd GED, Geurts R, Cannon SB, Udvardi MK, et al. (2011) The *Medicago* genome provides insight into the evolution of rhizobial symbioses. Nature 480: 520–524. doi: 10.1038/nature10625 PMID: 22089132

47. Hellsten U, Wright KM, Jenkins J, Shu S, Yuan Y, Wessler SR, et al. (2013) Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. Proc Natl Acad Sci USA 110: 19478–19482. doi: 10.1073/pnas.1319032110 PMID: 24225854

48. Kawahara Y, la Bastide de M, Hamilton JP, Kanamori H, McCombie WR, Ouyang S, et al. (2013) Improvement of the *Oryza sativa* Nipponbare reference genome using next generation sequence and optical map data. Rice 6: 4. doi: 10.1186/1471-2164-8-278 PMID: 24280374

49. Schmutz J, McClean PE, Mamidi S, Wu GA, Cannon SB, Grimwood J, et al. (2014) A reference genome for common bean and genome-wide analysis of dual domestications. Nat Genet 46: 707–713. doi: 10.1038/ng.3008 PMID: 24908249

50. Zimmer AD, Lang D, Buchta K, Rombauts S, Nishiyama T, Hasebe M, et al. (2013) Reannotation and extended community resources for the genome of the non-seed plant *Physcomitrella patens* provide insights into the evolution of plant gene structures and functions. BMC Genomics 14: 498. doi: 10.1016/S0168-9525(00)02093-X PMID: 23879659

51. The International Peach Genome Initiative (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. Nat Genet 45: 487–494. doi: 10.1038/ng.2586 PMID: 23525075

52. Tuskan GA, DiFazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, et al. (2006) The Genome of Black Cottonwood, *Populus trichocarpa* (Torr. & Gray). Science 313: 1596–1604. doi: 10.1126/science.1128691 PMID: 16973872

53. Chan AP, Crabtree J, Zhao Q, Lorenzi H, Orvis J, Puiu D, et al. (2010) Draft genome sequence of the oilseed species *Ricinus communis*. Nat Biotechnol 28: 951–956. doi: 10.1038/nbt.1674 PMID: 20729833

54. Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S, et al. (2012) Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. Nat Biotechnol 30: 549–554. doi: 10.1038/nbt.2195 PMID: 22580950

55. The Tomato Genome Consortium (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485: 635–641. doi: 10.1038/nature11119 PMID: 22660326

56. The Potato Genome Sequencing Consortium (2011) Genome sequence and analysis of the tuber crop potato. Nature 475: 189–195. doi: 10.1038/nature10158 PMID: 21743474

57. Wang W, Haberer G, Gundlach H, Gläßer C, Nussbaumer T, Luo MC, et al. (2014) The *Spirodela polyrhiza* genome reveals insights into its neotenous reduction fast growth and aquatic lifestyle. Nat Commun 5: 3311. doi: 10.1038/ncomms4311 PMID: 24548928

58. Motamayor JC, Mockaitis K, Schmutz J, Haiminen N, Livingstone D, Cornejo O, et al. (2013) The genome sequence of the most widely cultivated cacao type and its use to identify candidate genes regulating pod color. Genome Biol 14: R53. doi: 10.1186/gb-2013-14-6-r53 PMID: 23731509

59. The French—Italian Public Consortium for Grapevine Genome Characterization (2007) The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449: 463–467. doi: 10.1038/nature06148 PMID: 17721507

60. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, et al. (2009) The B73 Maize Genome: Complexity, Diversity, and Dynamics. Science 326: 1112–1115. doi: 10.1126/science.1178534 PMID: 19965430

61. Huang S, Ding J, Deng D, Tang W, Sun H, Liu D, et al. (2013) Draft genome of the kiwifruit *Actinidia chinensis*. Nat Commun 4: 2640. doi: 10.1038/ncomms3640 PMID: 24136039

62. Dohm JC, Minoche AE, Holtgräwe D, Capella-Gutiérrez S, Zakrzewski F, Tafer H, et al. (2014) The genome of the recently domesticated crop plant sugar beet (*Beta vulgaris*). Nature 505: 546–549. doi: 10.1038/nature12817 PMID: 24352233

63. Varshney RK, Chen W, Li Y, Bharti AK, Saxena RK, Schlueter JA, et al. (2011) Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. Nat Biotechnol 30: 83–89. doi: 10.1038/nbt.2022 PMID: 22057054

64. Qin C, Yu C, Shen Y, Fang X, Chen L, Min J, et al. (2014) Whole-genome sequencing of cultivated and wild peppers provides insights into Capsicum domestication and specialization. Proc Natl Acad Sci USA 111: 5135–5140. doi: 10.1073/pnas.1400975111 PMID: 24591624

65. Varshney RK, Song C, Saxena RK, Azam S, Yu S, Sharpe AG, et al. (2013) Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement. Nat Biotechnol 31: 240–246. doi: 10.1038/nbt.2491 PMID: 23354103

66. Sato S, Nakamura Y, Kaneko T, Asamizu E, Kato T, Nakao M, et al. (2008) Genome Structure of the Legume, *Lotus japonicus*. DNA Res 15: 227–239. doi: 10.1093/dnares/dsn008 PMID: 18511435

67. D'Hont A, Denoeud F, Aury J- M, Baurens F- C, Carreel F, Garsmeur O, et al. (2012) The banana (*Musa acuminata*) genome and the evolution of monocotyledonous plants. Nature 488: 213–217. doi: 10.1038/nature11241 PMID: 22801500

68. Chen J, Huang Q, Gao D, Wang J, Lang Y, Liu T, et al. (2013) Whole-genome sequencing of *Oryza brachyantha* reveals mechanisms underlying *Oryza* genome evolution. Nat Commun 4: 1–9. doi: 10.1038/ncomms2596

69. Al-Dous EK, George B, Al-Mahmoud ME, Al-Jaber MY, Wang H, Salameh YM, et al. (2011) *De novo* genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). Nat Biotechnol 29: 521–527. doi: 10.1038/nbt.1860 PMID: 21623354

70. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin Y-C, Scofield DG, et al. (2013) The Norway spruce genome sequence and conifer genome evolution. Nature 497: 579–584. doi: 10.1038/nature12211 PMID: 23698360

71. Wegrzyn JL, Liechty JD, Stevens KA, Wu L-S, Loopstra CA, Vasquez-Gross HA, et al. (2014) Unique Features of the Loblolly Pine (*Pinus taeda* L.) Megagenome Revealed Through Sequence Annotation. Genetics 196: 891–909. doi: 10.1534/genetics.113.159996 PMID: 24653211

72. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2010) GenBank. Nucleic Acids Res 38: D46–D51. doi: 10.1093/nar/gkp1024 PMID: 19910366

73. Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, et al. (2009) Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25: 1422–1423. doi: 10.1093/bioinformatics/btp163 PMID: 19304878

74. Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Research 39: W29–W37. doi: 10.1093/nar/gkr367 PMID: 21593126

75. Magrane M, UniProt Consortium (2011) UniProt Knowledgebase: a hub of integrated protein data. Database 2011: bar009–bar009. doi: 10.1093/database/bar009 PMID: 21447597

76. Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 48: 443–453. PMID: 5420325

77. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215: 403–410. doi: 10.1016/S0022-2836(05)80360-2 PMID: 2231712

78. Dosztányi Z, Csizmók V, Tompa P, Simon I (2005) The Pairwise Energy Content Estimated from Amino Acid Composition Discriminates between Folded and Intrinsically Unstructured Proteins. J Mol Biol 347: 827–839. PMID: 15769473

79. Millman KJ, Aivazis M (2011) Python for Scientists and Engineers. Comput Sci Eng 13: 9–12. doi: 10.1109/MCSE.2011.36

80. Oliphant TE (2007) Python for Scientific Computing. Comput Sci Eng 9: 10–20. doi: 10.1109/MCSE.2007.58

81. Kans J (2013) Entrez Direct: E-utilities on the UNIX Command Line. In: Entrez Programming Utilities Help (Internet). Bethesda (MD): National Center for Biotechnology Information (US); 2010-. Available: http://www.ncbi.nlm.nih.gov/books/NBK179288/. Accessed 11 January 2015.

82. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37: W202–W208. doi: 10.1093/nar/gkp335 PMID: 19458158

83. Fauteux F, Blanchette M, Strömvik MV (2008) Seeder: discriminative seeding DNA motif discovery. Bioinformatics 24: 2303–2307. doi: 10.1093/bioinformatics/btn444 PMID: 18718942

84.   Pavesi G, Mauri G, Pesole G (2001) An algorithm for finding signals of unknown length in DNA sequences. Bioinformatics 17: S207–S214. doi: 10.1093/bioinformatics/17.suppl_1.S207 PMID: 11473011

85.   Mahony S, Benos PV (2007) STAMP: a web tool for exploring DNA-binding motif similarities. Nucleic Acids Res 35: W253–W258. doi: 10.1093/nar/gkm272 PMID: 17478497

86.   Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: A Sequence Logo Generator. Genome Res 14: 1188–1190. doi: 10.1101/gr.849004 PMID: 15173120

87.   Close TJ (1996) Dehydrins: emergence of a biochemical role of a family of plant dehydration proteins. Physiol Plant 97: 795–803.

88.   Tunnacliffe A, Wise MJ (2007) The continuing conundrum of the LEA proteins. Naturwissenschaften 94: 791–812–812. doi: 10.1007/s00114-007-0254-y PMID: 17479232

89.   Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105–132. doi: 10.1016/0022-2836(82)90515-0 PMID: 7108955

90.   Fauteux F, Strömvik MV (2009) Seed storage protein gene promoters contain conserved DNA motifs in Brassicaceae, Fabaceae and Poaceae. BMC Plant Biol 9: 126. doi: 10.1186/1471-2229-9-126 PMID: 19843335

91.   Rorat T, Grygorowicz WJ, Irzykowski W, Rey P (2003) Expression of KS-type dehydrins is primarily regulated by factors related to organ type and leaf developmental stage during vegetative growth. Planta 218: 878–885. doi: 10.1007/s00425-003-1171-8 PMID: 14685858

92.   Tommasini L, Svensson J, Rodriguez E, Wahid A, Malatrasi M, Kato K, et al. (2008) Dehydrin gene expression provides an indicator of low temperature and drought stress: transcriptome-based analysis of barley (*Hordeum vulgare* L.). Funct Integr Genomic 8: 387–405–405. doi: 10.1007/s10142-008-0081-z PMID: 18512091

93.   Hara M, Shinoda Y, Kubo M, Kashima D, Takahashi I, Kato T, et al. (2011) Biochemical characterization of the *Arabidopsis* KS-type dehydrin protein, whose gene expression is constitutively abundant rather than stress dependent. Acta Physiol Plant 33: 2103–2116–2116. doi: 10.1007/s11738-011-0749-1

94.   Martínez-Hernández A, López-Ochoa L, Argüello-Astorga G, Herrera-Estrella L (2002) Functional properties and regulatory complexity of a minimal RBCS light-responsive unit activated by phytochrome, cryptochrome, and plastid signals. Plant Physiol 128: 1223–1233. doi: 10.1104/pp.010678 PMID: 11950971

95.   Degenhardt J, Tobin EM (1996) A DNA binding activity for one of two closely defined phytochrome regulatory elements in an *Lhcb* promoter is more abundant in etiolated than in green plants. Plant Cell 8: 31–41. PMID: 8597658

96.   Tatematsu K, Ward S, Leyser O, Kamiya Y, Nambara E (2005) Identification of cis-elements that regulate gene expression during initiation of axillary bud outgrowth in Arabidopsis. Plant Physiol 138: 757–766. PMID: 15908603

97.   Kaplan F, Kopka J, Haskell DW, Zhao W, Schiller KC, Gatzke N, et al. (2004) Exploring the temperature-stress metabolome of Arabidopsis. Plant Physiol 136: 4159–4168. doi: 10.1104/pp.104.052142 PMID: 15557093

98.   Hoekstra FA, Golovina EA, Buitink J (2001) Mechanisms of plant desiccation tolerance. Trends Plant Sci 6: 431–438. doi: 10.1016/S1360-1385(01)02052-0 PMID: 11544133

99.   Ohno R, Takumi S, Nakamura C (2003) Kinetics of transcript and protein accumulation of a low-molecular-weight wheat LEA D-11 dehydrin in response to low temperature. J Plant Physiol 160: 193–200. doi: 10.1078/0176-1617-00925 PMID: 12685035

100.   Šunderlíková V, Salaj J, Kopecky D, Salaj T, Wilhem E, Matušíková I (2009) Dehydrin genes and their expression in recalcitrant oak (*Quercus robur*) embryos. Plant Cell Rep 28: 1011–1021. doi: 10.1007/s00299-009-0710-6 PMID: 19466427

101.   Kosugi S, Suzuka I, Ohashi Y (1995) Two of three promoter elements identified in a rice gene for proliferating cell nuclear antigen are essential for meristematic tissue-specific expression. Plant J 7: 877–886. PMID: 7599648

102.   Kasuga M, Liu Q, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1999) Improving plant drought, salt, and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Nat Biotechnol 17: 287–291. doi: 10.1038/7036 PMID: 10096298

103.   Nordin K, Vahala T, Palva ET (1993) Differential expression of two related, low-temperature-induced genes in *Arabidopsis thaliana* (L.) Heynh. Plant Mol Biol 21: 641–653. doi: 10.1007/BF00014547 PMID: 8448363

104.   Busk PK, Pagès M (1998) Regulation of abscisic acid-induced transcription. Plant Mol Biol 37: 425–435–435. doi: 10.1023/A:1006058700720 PMID: 9617810

105. Deng Z, Pang Y, Kong W, Chen Z, Wang X, Liu X, et al. (2013) A novel ABA-dependent dehydrin *ERD10* gene from *Brassica napus*. Mitochondr DNA 16: 28–35. doi: 10.1080/10425170500040180

106. Ezcurra I, Ellerström M, Wycliffe P, Stålberg K, Rask L (1999) Interaction between composite elements in the *napA* promoter: both the B-box ABA-responsive complex and the RY/G complex are necessary for seed-specific expression. Plant Mol Biol 40: 699–709. PMID: 10480393

107. Busk PK, Pagès M (1997) Protein binding to the abscisic acid-responsive element is independent of VIVIPAROUS1 in vivo. Plant Cell 9: 2261–2270. doi: 10.1105/tpc.9.12.2261 PMID: 11407411

108. Ono A, Izawa T, Chua N-H, Shimamoto K (1996) The *rab16B* promoter of rice contains two distinct abscisic acid-responsive elements. Plant Physiol. 112, 483–491. PMID: 8883374

109. Straub PF, Shen Q, Ho TD (1994) Structure and promoter analysis of an ABA- and stress-regulated barley gene, HVA1. Plant Mol Biol 26: 617–630. PMID: 7948917

110. Qin F, Sakuma Y, Li J, Liu Q, Li Y-Q, Shinozaki K, et al. (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. doi: 10.1093/pcp/pch118 PMID: 15356330

111. Skinner JS, Zitzewitz J, Sz cs P, Marquez-Cedillo L, Filichkin T, Amundsen K, et al. (2005) Structural, Functional, and Phylogenetic Characterization of a Large *CBF* Gene Family in Barley. Plant Mol Biol 59: 533–551. doi: 10.1007/s11103-005-2498-2 PMID: 16244905

112. Suzuki M, Ketterling MG, McCarty DR (2005) Quantitative statistical analysis of *cis*-regulatory sequences in ABA/VP1- and CBF/DREB1-regulated genes of Arabidopsis. Plant Physiol 139: 437–447. doi: 10.1104/pp.104.058412 PMID: 16113229

113. Giuliano G, Pichersky E, Malik VS, Timko MP, Scolnik PA, Cashmore AR (1988) An evolutionarily conserved protein binding sequence upstream of a plant light-regulated gene. Proc Natl Acad Sci USA 85: 7089–7093. PMID: 2902624

114. Donald RGK, Cashmore AR (1990) Mutation of either G box or I box sequences profoundly affects expression from the *Arabidopsis rbcS-1A* promoter. EMBO J 9: 1717. PMID: 2347304

115. Ngai N, Tsai F-Y, Coruzzi G (1997) Light-induced transcriptional repression of the pea *AS1* gene: identification of *cis*-elements and transfactors. Plant J 12: 1021–1034. doi: 10.1046/j.1365-313X.1997.12051021.x PMID: 9418044

116. Stockinger EJ, Gilmour SJ, Thomashow MF (1997) *Arabidopsis thaliana CBF1* encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. Proc Natl Acad Sci USA 94: 1035–1040. PMID: 9023378

117. Fowler SG, Cook D, Thomashow MF (2005) Low Temperature Induction of Arabidopsis *CBF1*, *2*, and *3* Is Gated by the Circadian Clock. Plant Physiol 137: 961–968. doi: 10.1104/pp.104.900140 PMID: 15728337

118. Tóth R, Kevei E, Hall A, Millar AJ, Nagy F, Kozma-Bognár L (2001) Circadian clock-regulated expression of phytochrome and cryptochrome genes in Arabidopsis. Plant Physiol 127: 1607–1616. PMID: 11743105

119. Mikkelsen MD, Thomashow MF (2009) A role for circadian evening elements in cold-regulated gene expression in Arabidopsis. Plant J 60: 328–339. doi: 10.1111/j.1365-313X.2009.03957.x PMID: 19566593

120. Rawat R, Xu Z-F, Yao K-M, Chye M-L (2005) Identification of cis-elements for ethylene and circadian regulation of the *Solanum melongena* gene encoding cysteine proteinase. Plant Mol Biol 57: 629–643. doi: 10.1007/s11103-005-0954-7 PMID: 15988560

121. Kreps J, Budworth P, Goff S, Wang R (2003) Identification of putative plant cold responsive regulatory elements by gene expression profiling and a pattern enumeration algorithm. Plant Biotechnol J 1: 345–352. doi: 10.1046/j.1467-7652.2003.00032.x PMID: 17166133

122. Hong JC, Cheong YH, Nagao RT, Bahk JD, Key JL, Cho MJ (1995) Isolation of two soybean G-box binding factors which interact with a G-box sequence of an auxin-responsive gene. Plant J 8: 199–211. PMID: 7670504

123. Xu N, Hagen G, Guilfoyle T (1997) Multiple auxin response modules in the soybean *SAUR 15A* promoter. Plant Sci 126: 193–201. doi: 10.1016/S0168-9452(97)00110-6

124. Lee B-H, Henderson DA, Zhu J-K (2005) The Arabidopsis cold-responsive transcriptome and its regulation by ICE1. Plant Cell 17: 3155–3175. doi: 10.1105/tpc.105.035568 PMID: 16214899

125. Shih M-D, Hoekstra FA, Hsing Y-IC (2008) Late Embryogenesis Abundant Proteins. In: Delseny J-CKAM, editor. Advances in Botanical Research. Advances in Botanical Research. Academic Press, Vol. Volume 48. pp. 211–255.

126. Archambault A, Strömvik MV (2011) The Y-segment of novel cold dehydrin genes is conserved and codons in the PR-10 genes are under positive selection in *Oxytropis* (Fabaceae) from contrasting climates. Mol Genet Genomics 287: 123–142. doi: 10.1007/s00438-011-0664-6 PMID: 22183143

127. Sarnighausen E, Karlson DT, Zeng Y, Goldsbrough PB, Raghothama KG, Ashworth EN (2004) Characterization of a Novel YnSKn Class of Dehydrin-Like cDNAs from Cold Acclimated Red-Osier Dogwood (*Cornus sericea* L.) Xylem. J Crop Improv 10: 17–35. doi: 10.1300/J411v10n01_03

128. Garcia-Bañuelos ML, Gardea AA, Winzerling JJ, Vazquez-Moreno L (2009) Characterization of a Midwinter-Expressed Dehydrin (DHN) Gene from Apple Trees (*Malus domestica*). Plant Mol Biol Rep 27: 476–487. doi: 10.1007/s11105-009-0110-7

129. Welling A, Rinne P, Vihera-Aarnio A, Kontunen-Soppela S, Heino P, Palva TE (2004) Photoperiod and temperature differentially regulate the expression of two dehydrin genes during overwintering of birch (*Betula pubescens* Ehrh.). J Exp Bot 55: 507–516. doi: 10.1093/jxb/erh045 PMID: 14739271

130. Wisniewski ME, Bassett CL, Renaut J, Farrell R, Tworkoski T, Artlip TS (2006) Differential regulation of two dehydrin genes from peach (*Prunus persica*) by photoperiod, low temperature and water deficit. Tree Physiol 26: 575–584. PMID: 16452071

131. Finch-Savage WE, Pramanik SK, Bewley JD (1994) The expression of dehydrin proteins in desiccation-sensitive (recalcitrant) seeds of temperate trees. Planta 193: 478–485. doi: 10.1007/BF02411551

132. Delahaie J, Hundertmark M, Bove J, Leprince O, Rogniaux H, Buitink J (2013) LEA polypeptide profiling of recalcitrant and orthodox legume seeds reveals *ABI3*-regulated LEA protein abundance linked to desiccation tolerance. J Exp Bot. doi: 10.1093/jxb/ert274