ISSN 1993-6842 (on-line); ISSN 0233-7657 (print)
Biopolymers and Cell. 2023. Vol. 39. N 2. P 110–130
http://dx.doi.org/10.7124/bc.000A8E

UDC 577.21+582.542.1

Prediction and analysis of stress-inducible ICE transcription factors in *Deschampsia antarctica*

O. M. Bublyk, I. O. Andreev, V. A. Kunakh
Institute of Molecular Biology and Genetics, NAS of Ukraine
150, Akademika Zabolotnoho Str., Kyiv, Ukraine, 03143
o.m.bublyk@imbg.org.ua

Aim. To study molecular genetic aspects of stress tolerance in the extremophile plant *D. antarctica* by exploring ICE transcription factors involved in cold stress response, regulation of stomatal development and flowering time. **Methods.** The *DaICE* genes were assembled from SRA reads homologous to orthologous genes used as a reference. The predicted genes and proteins were analyzed with bioinformatics methods. Gene expression was analyzed using the RNA-seq data. **Results.** Two variants of the *DaICE1* gene were found: a variant that encodes a full-length protein and a variant with a large deletion, which led to the loss of the N-terminus and part of the bHLH domain in the protein product. Two variants of CDS differed by four single-nucleotide polymorphisms, three of which were synonymous, with more dissimilar promoter sequences were found for the *DaICE2*. The organization of promoters and the intron-exon structure of the genes were studied. For the predicted TF proteins, physico-chemical properties were determined, nuclear localization and the possibility of post-translational modifications were shown, phylogenetic analysis was carried out, conservative motifs were found, 3D structure was predicted. Gene expression was detected in most or all plant tissues and increased expression of *DaICE1* variants was shown in the plants from natural conditions compared to laboratory ones. **Conclusions.** DaICE1 and DaICE2 are the candidate regulators of the CBF gene expression under cold stress in *D. antarctica*.

Keywords: Antarctic plants, transcription factors, *in silico* prediction, abiotic stress, resistance.

Introduction

The significant negative impact of low temperatures on the productivity of agricultural crops, including cereal crops of the Poaceae, highlights the relevance of studying the molecular mechanisms of cold response in plants. *Deschampsia antarctica* È. Desv. is a useful object for such studies as it is the only grass species that has adapted to the extreme envi-

ronmental conditions of the maritime Antarctic characterized by low temperatures, strong winds, low soil water availability, and intense solar radiation [1]. Due to its hardiness, *D. antarctica* can serve as a source of genes to enhance the cold resistance of economically valuable cereals through genetic modification [2–4]. However, the study of the molecular genetic mechanisms of stress response in this species has only just begun. Specifically, Lee *et al.* [5, 6] investigated general changes in the expression of different families of transcription factors (TFs) in *D. antarctica* compared to other grasses and *A. thaliana* under abiotic stress conditions such as low temperature, drought, and salinity. DaCBF4 and DaCBF7 (C-repeat binding factor) are stress-inducible TFs that have been extensively studied [4, 5]. Furthermore, in our previous study [7], we identified through whole-genome analysis and described 17 TFs from the CBF/DREB1 (dehydration responsive element-binding protein 1) and DREB4 groups of *D. antarctica*.

CBF/DREB1 is a major group of TFs responsible for the induction of cold tolerance [8]. These genes are activated in response to low temperatures by TFs known as inducers of CBF expression (ICE), which belong to the MYC-like bHLH (myelocytomatosis-like basic helix-loop-helix) proteins and bind to the E-box cis-element (CANNTG) [9]. The ICE–CBF–COR (cold-responsive) pathway is considered the main signaling pathway of cold stress response in plants [10]. ICE TFs were first identified in *A. thaliana*: it was shown that ICE1 activates *AtCBF3* [9] and ICE2 activates *AtCBF1* genes [11]. Similarly, Badawi *et al.* [12] demonstrated that TaICE87 and TaICE41 TFs, which respectively belong to the ICE1 and ICE2 groups, activate the transcription of *CBF* genes in wheat. Homologous *ICE* genes have now been identified in many other plant species, including grasses [10]. There are a number of reports demonstrating that overexpression of the *ICE* genes in transformed plants can improve the cold tolerance [11–14] as well as drought and salt tolerance [15, 16].

In addition to the cold response, ICE TFs are also involved in regulation of several physiological processes in plants. In particular, ICE1 TFs have been shown to be involved in stomatal development [17], inhibition of flowering under cold stress conditions [18], regulation of primary seed dormancy [19], seed germination [20], endosperm development [21] and male fertility [22]. The studies of the role of ICE2 in various processes showed that this TF is also multifunctional, affecting the development of stomata and flowering time [17, 23].

ICE TFs are expressed constitutively and are present in an inactive state at normal temperatures. When plants are exposed to low temperatures, ICE1 are activated through sumoylation by the SIZ1 (SAP and Miz 1) protein known as SUMO E3 ligase (small ubiquitin-related modifier) [24]. The activation process is opposed by the E3 RING finger ligase HOS1 (high expression of osmotically responsive gene 1) that mediates ubiquitination and degradation of ICE1 [25]. In addition, ICE1 is phosphorylated by the cold-activated OST1 (OPEN STOMATA 1) protein kinase, which weakens the interaction between ICE1 and HOS1 and increases ICE1 stability [26]. OsICE1 stability is also regulated by OsMPK3 (MAP KINASE 3) through phosphorylation in response to cold stress [27].

The aim of this study was to continue our previous study [7] on the cold resistance mechanisms of *D. antarctica* by exploring the ICE–CBF signaling pathway. Specifically, we conducted *in silico* prediction of the genes encoding ICE TFs of *D. antarctica*, analyzed the intron-exon structure of the genes and organization of their promoters as well as characterized the encoded proteins. The characterization of the ICE TFs included the determination of physicochemical properties, subcellular localization, sites of post-translational modifications, phylogenetic analysis, search for conservative motifs, 3D structure prediction, and comparison of their expression in different tissues of plants grown under different conditions, based on RNA-seq analysis.

Materials and Methods

The fragments of *D. antarctica ICE* genes were searched for using the BLAST program [28] in the SRA (sequence read archive) bioinformatics database containing the raw data of *D. antarctica* genome (SRX465632) and transcriptome (SRX465633, SRX5305101 — SRX5305108) sequencing, which are available in GenBank (https://www.ncbi.nlm.nih.gov/genbank/). The sequences of orthologous genes from other grasses available in GenBank were used as references. The assembly of complete sequences of *DaICE* genes from the identified fragments was performed using the MEGA X program [29]. To perform a whole-genome analysis of the putative ICE genes, the sequences with high homology to assembled genes were also searched. The identified genes and translated proteins were confirmed to belong to ICE group by search for homological sequences in GenBank.

The intron-exon structure of the assembled *DaICE* genes was visualized using GSDS (Gene Structure Display Server) (http://gsds.gao-lab.org) [30].

The alignment of *D. antarctica* ICE proteins using the Muscle algorithm [31] and their phylogenetic analysis involving homologues sequences from other grasses and *A. thaliana* using the Maximum Likelihood method and the JTT matrix-based model [32] were performed in the MEGA X software with 1000 bootstrap replicates [29]. Alignment was visualized using GeneDoc ver. 2.7 [33]. The presence of ICE group-specific bHLH and ACT-like domains (the latter is named after the three proteins that contain it: aspartate kinase, chorismate mutase, and TyrA) was confirmed in the CDART web-based tool (https://www.ncbi.nlm.nih.gov/Structure/lexington/lexington.cgi) [34].

Physicochemical properties of DaICE proteins, including molecular weight (Mw), theoretical isoelectric point (pI), and instability index, were calculated using the ProtParam tool (https://web.expasy.org/protparam/) [35]. The Plant-mPLoc server [36] was used to predict subcellular localization, and the nuclear localization signals (NLS) were identified using the INSP (Identification Nucleus Signal Peptide) method (http://www.csbio.sjtu.edu.cn/bioinf/INSP/) [37].

Post-translational modifications were predicted using MusiteDeep (https://musite.net) [38] with a threshold of 0.5.

Conserved motifs in ICE protein sequences were detected using MEME Suite ver. 5.5.1 (Multiple Em for Motif Elicitation) (https://meme-suite.org/) [39] with the following parameters: maximum number of motifs to

search — 12; any number of repetitions per motif; motif length to search — 6–100.

The secondary structure of DaICE proteins was predicted using the JPred4: a protein secondary structure prediction server (https://www.compbio.dundee.ac.uk/jpred/) [40]. The three-dimensional structure of proteins was predicted using the Phyre2 (Protein Homology/Analogy Recognition Engine ver. 2.0) web portal in the "intensive" mode (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index) [41]. The structural changes caused by amino acid substitution in the two DaICE2 variants were calculated on the Missense3D server (http://missense3d.bc.ic.ac.uk/missense3d) [42].

The *cis*-elements in promoter sequences of *DaICE* genes were searched for using the PlantCare online tools [43]. Pairwise distances according to the Maximum Composite Likelihood method between promoter sequences were calculated using the MEGA X program [29], taking into account transitions and transversions.

To assess the expression level of *DaICE* genes, the number of reads homologous to them was counted in the RNA sequencing data archives (SRX5305101 — SRX5305108). These archives contain transcriptomic data of roots, flowers, crowns and leaves of *D. antarctica* wild growing-plants collected on the Barton Peninsula (King George Island), near the Korean Antarctic Station King Sejong in January 2013 and plants grown in the growth cabinet. The obtained values were normalized using the RPKM formula (Reads per kilobase of transcript per million mapped reads) $R = 10^9C/NL$, where C is the number of homologous reads found, N is the total number of reads in the archive, and L is the gene length in base pairs [44]. The gene fragments of 300 bp from the regions that differed in two *DaICE1* variants and two *DaICE2* variants were taken for analysis. The stability of gene expression was evaluated using the BestKeeper tool available at Reffinder (https://www.heartcure.com.au/reffinder/) [45].

Results and Discussion

As a result of the search conducted in the genome and transcriptome sequencing databases of *D. antarctica* using the sequences of known orthologous genes *ICE1* and *ICE2* from other grasses as the reference, we have assembled the genes of *DaICE1* and *DaICE2* TFs. The assembled sequences were deposited in the Genbank database under accession numbers OR091061 to OR091064.

For the *DaICE1* gene, based on the genome and transcriptome sequencing data, two variants were assembled: a normal one encoding a full-length protein, and a variant with a large deletion of 914 nucleotides covering parts of exon 1 and intron 1, which we named *DaICE1sh* (Fig. 1). In this variant, intron 1 is not excised, thus resulting in the shift of initial reading frame, which prevents the translation of the full-length normal protein. However, within this transcript, there is another open reading frame encoding a part of the DaICE1 protein. It has a start codon at 1–3 bp of exon 2 and potentially can be translated into a protein.

For the *DaICE2* gene, two very similar coding sequence variants were assembled, which differed by only four single nucleotide polymorphisms (SNPs). Three of them were synonymous, and one led to a substitution of alanine (A) by valine (V) at position 109. These

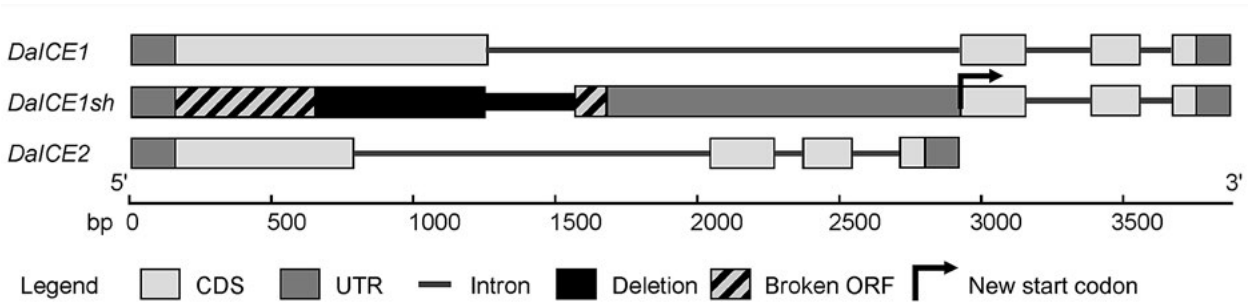

Fig. 1. Intron-exon structure of *ICE* genes of *D. antarctica*.

amino acids have similar characteristics, both are non-polar, aliphatic monoamino monocarboxylic acids. According to the Missense3D server modeling data [42], this substitution also did not lead to a change in the 3D structure of the protein. These two gene variants may be alleles or duplicated loci of the gene.

However, the promoter regions upstream of these almost identical *DaICE2* coding sequence variants differed significantly. In particular, they contained numerous substitutions, insertions, and deletions. The pairwise distances between the two *DaICE2* promoter variants, calculated by the Maximum Composite Likelihood method (transitions and transversions were taken into account) in the MEGA X program, were 0.054. For comparison, we added to the analysis the promoter of the *LrICE1-like* gene (XP_047058470.1), which was found to be the closest to *DaICE2* according to the BLAST search in GenBank. The corresponding sequence upstream of the *LrICE1-like* gene was found on chromosome 6 of *Lolium rigidum* (CM040113.1). The pairwise distances between the promoter of this gene and variants 1 and 2 of the *DaICE2* promoter were 0.242 and 0.234, respectively. Therefore, the differences between the variants of the *DaICE2* promoter were on average 4.4 times lower than the interspecific differences between the phylogenetically closest species.

Analysis of the intron-exon structure of the genes, carried out by comparing the assembled sequences from genomic DNA and mRNA archives, revealed that *DaICE1* and *DaICE2* each have three introns and four exons (Fig. 1). According to the literature data, orthologous genes of other species have the same structure. Their primary transcripts undergo processing, where three introns are removed [46, 47]. The borders of all predicted *ICE1* and *ICE2* introns corresponded to the GT-AG rule.

Analysis of the predicted DaICE proteins revealed that DaICE1 has a specific sequence for the ICE1 TFs, GAQPTLFQKRA, which has been described previously [12]. Additionally, an ICE-specific sequence, KMDRASILGDAIEYLKELL, was found in DaICE1 [48], and a variant of this sequence, KMDRASILGDAIDYLKELL, was found in DaICE2 (Fig. 2).

The predicted DaICE proteins were found to possess typical bHLH and ACT-like domains (Fig. 2), the presence of which was confirmed using the CDART web-based

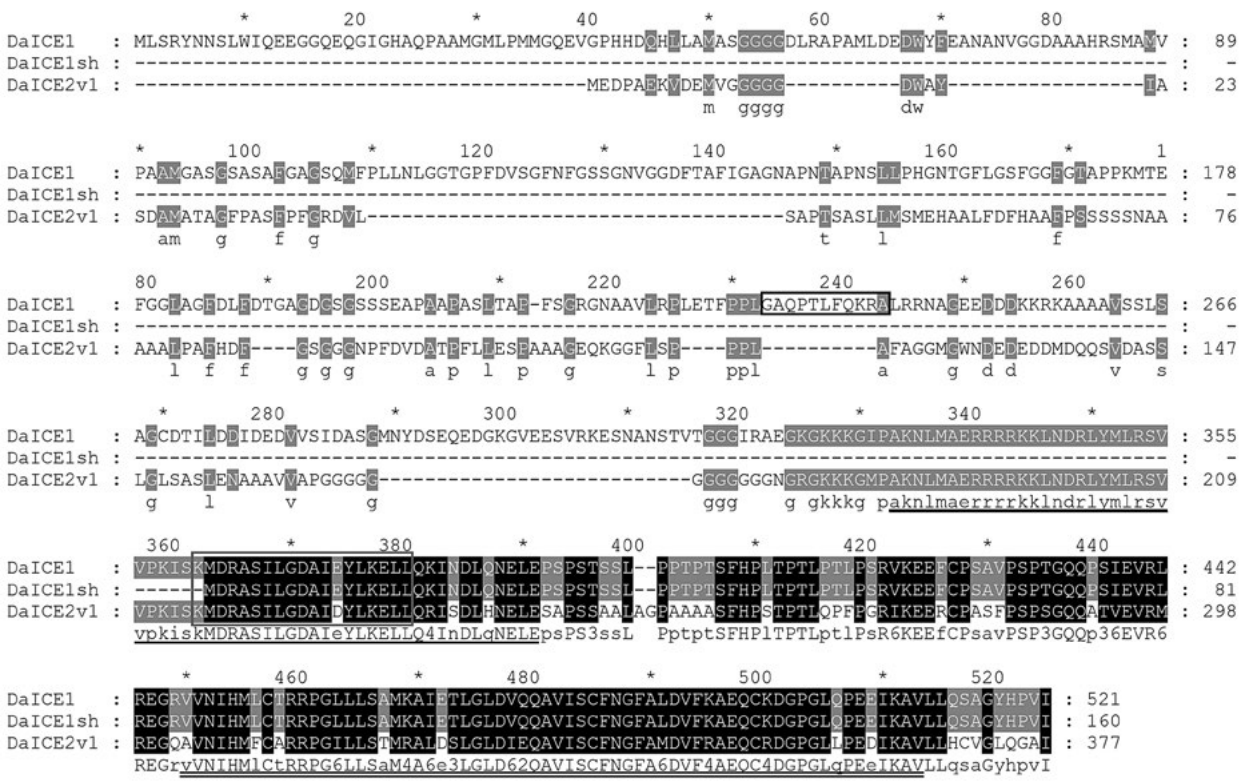

Fig. 2. Alignment of putative amino acid sequences of ICE proteins of *D. antarctica*. Conserved amino acid residues among all or most of the sequences are highlighted in black and gray, respectively. Characteristic sequences are boxed. The bHLH and ACT-like domains are single and double underlined, respectively.

tool [34]. The bHLH domain is responsible for DNA binding and homo- or heterodimerization of the ICE TF. This domain participates in the formation of ICE homodimers, which are involved in cold response [11] and regulation of flowering time [18], and heterodimers with other TFs, which are involved in the regulation of endosperm breakdown [21] and seed germination [20].

The ACT-like domain of ICE1 (also known as SCREAM) is responsible for the formation of heterodimers with other bHLH TFs. It plays an important role in dimerization partner selection and in switching from proliferation to differentiation during stomatal development. Stomatal differentiation is regulated by the sequential actions of three bHLH proteins: SPEECHLESS (SPCH), MUTE, and FAMA, responsible for initiation and proliferation, commitment, and terminal differentiation, respectively. The ICE1 TF heterodimerizes with SPCH/MUTE/FAMA and participates in the regulation of stomatal differentiation via switching its partners [49].

Unlike *DaICE1*, the open reading frame of the short variant of the *DaICE1* gene, *DaICE1sh*, encodes a truncated protein that starts at two amino acids upstream of the beginning of the

last helix of the bHLH domain [9]. A search for the amino acid sequences homologous to DaICE1 and DaICE1sh in GenBank revealed a similar pair of proteins in sorghum (*Sorghum bicolor*): the normal XP_002459022.1 and the truncated XP_002448735.1, which is of the same length as DaICE1sh. These proteins are encoded by two different genes located in genomic loci LOC8059393 and LOC8082581 on chromosomes 3 and 6, respectively. These TFs belong to a different group, namely ICE2, but have similar functions, particularly in the cold response and stomatal development.

As a result of the deletion, DaICE1sh lost a part of the bHLH domain and the ability to homodimerize and bind to DNA, but retained the ACT-like domain responsible for protein heterodimerization. Such truncated proteins have been described in the literature for other genes and referred to as microProteins [50] or small interfering peptides (siPEP) [51]. They can be formed by deletions from functional TFs, as in our case, or by alternative splicing [51]. MicroProteins form non-functional dimers and act through the dominant-negative suppression of protein complex function. These proteins are quite common; they are involved in the regulation of a number of physiological processes, occur in animals and plants [50], and have been found in almost all TF families [52].

The potential protein product of *DaICE1sh* may be involved in the negative regulation of stomatal differentiation processes by forming non-functional dimers and blocking partner proteins. It was shown that the isolated ACT-like domain of SCRM TF (also known as ICE1) can stably heterodimerize with the same domain of SPCH and MUTE [49]. Therefore, it is likely that the truncated protein DaICE1sh may function in a similar way. The existence of such a truncated protein in sorghum, a highly drought-resistant species, may theoretically indicate the adaptiveness of this mutation, as the density, size, morphology, and functionality of stomata play an important role in the control of water loss and affect the level of drought resistance of a plant [53, 54].

Table 1. **Characteristics of putative *D. antarctica* ICE proteins: physicochemical properties, subcellular localization, and nuclear localization signals**

Name	Number of amino acids	Molecular weight, Mw (kDa)	Theoretical isoelectric point, pI	Instability index*	Localization	Nuclear localization signals, score
DaICE1	521	54,67	5,36	59,21	Nucleus	GAQPTLFQKRALRRNAGEEDDD-KKRK, 0,94 KGKKKGIPAKNL-MAERRRRKKLNDRLYMLRSVVP-KISKMD, 0,95
DaICE1sh	160	17,44	5,36	75,75	Nucleus	VRLREGRVVNI, 0,88
DaICE2v1	377	39,18	5,02	61,42	Nucleus	KNLMAERRRRKKLNDRLYMLRSV-VPKISKMD, 0,96
DaICE2v2	377	39,20	5,02	61,20	Nucleus	KNLMAERRRRKKLNDRLYMLRSV-VPKISKMD, 0,96

Note: * — the value of instability index above 40 predicts that the protein may be unstable.

The characteristics of putative *D. antarctica* ICE proteins are presented in Table 1. The number of amino acid residues ranged from 160 to 521. The estimated molecular weight ranged from 17.44 to 54.67 kDa. The theoretical isoelectric points were 5.02 and 5.36, and the instability index ranged from 59.21 to 75.75, with all proteins being predicted as unstable. Nuclear localization was predicted for all proteins, followed by the identification of nuclear localization signals. The latter is consistent with the literature data reporting that ICE proteins of various species are localized within the nucleus, which is a necessary condition to perform regulatory functions in the transcription process [9, 55–59].

We performed a phylogenetic analysis of the predicted DaICE proteins and ICE proteins of other grass species and *A. thaliana* retrieved from the GenBank database, a total of 12 species. Prior to the analysis, we searched the GenBank Genome database for each species to verify the number of *ICE* genes. In most species, including *A. thaliana, B. distachyon, H. vulgare, L. rigidum, O. brachyantha, O. sativa,* and *S. italic*, two *ICE* genes were found, one *ICE1* gene and one *ICE2* gene. Three *ICE* genes were found in the rest of the species, with *Ae. tauschii, T. aestivum*, and *Z. mays* having two *ICE1* genes and one *ICE2* gene, and *S. bicolor* having one *ICE1* gene and two *ICE2* genes. Figure 3 shows the numbers of chromosomes where the identified *ICE* genes are located in different species. According to the literature, most species typically have one or two *ICE* genes, with the exception of apple and grape, which have four genes each [46].

The phylogenetic analysis confirmed that the predicted proteins belong to the ICE group. Two clusters of the dendrogram included ICE1 and ICE2 orthologs from grasses, respectively, whereas the third cluster was formed by two *A. thaliana* ICE proteins (Fig. 3). Within the first two clusters, the grouping generally corresponded to the phylogenetic relationships among the species, belonging to the three subfamilies of the Poaceae family: Pooideae (*D. antarctica, H. vulgare, Ae. tauschii, L. rigidum, T. aestivum, B. distachyon*), Oryzoideae (*O. brachyantha, O. sativa*), and Panicoideae (*S. italica, Z. mays, S. bicolor*). All bootstrap values exceeded 50 %. DaICE1 had the highest homology with *L. rigidum* ICE1-like (XP_047088968.1), with a sequence identity of 88 % according to the BLAST analysis. DaICE2 had the highest sequence homology with *L. rigidum* ICE1-like (XP_047058470.1), with a sequence identity of 85 %. The coverage was 100 % in both cases.

Conserved motifs in predicted DaICE protein sequences were searched using the MEME Suite software [39]. Conserved motifs are often associated with biological functions, such as transcriptional activity, protein-protein interaction, and nuclear localization. Their presence can serve as a basis for protein classification. On the other hand, they are also related to evolutionary history of the species, and thus can be used in phylogenetic analysis [60].

In our analysis, we limited the number of displayed motifs to twelve, as this was the minimum number that allowed us to identify a motif common to most of the proteins from the ICE2 group (Fig. 4). The number of sites contributing to the construction of the motif varied from 7 to 36 and the length of found motifs ranged from 11 to 68 a.a. Motif 1 that contained

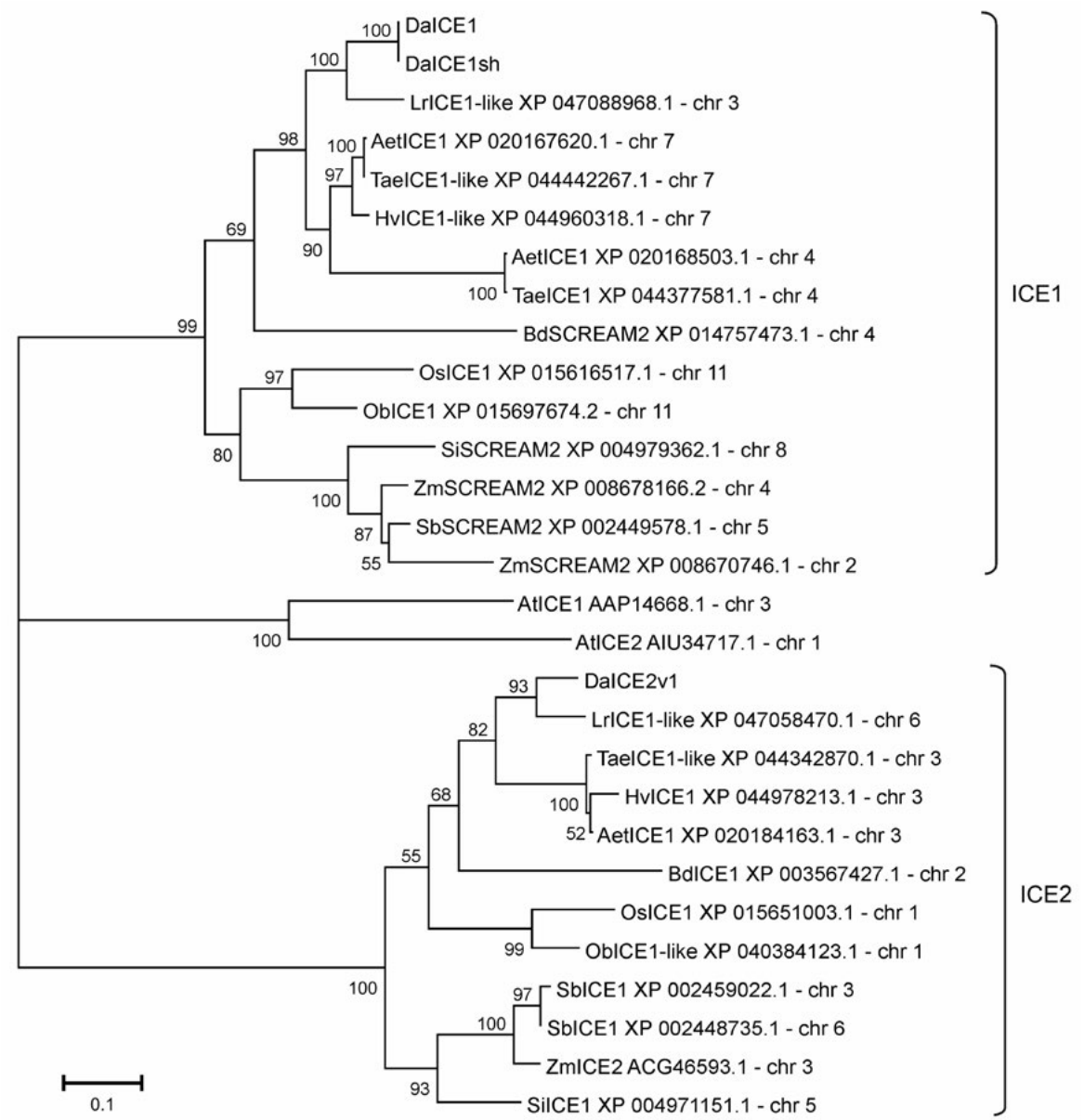

Fig. 3. Phylogenetic analysis of 29 ICE amino acid sequences of grasses and *A. thaliana*. *D. antarctica* predicted protein sequences and sequences from the GenBank database were used (the name, GenBank identification number, and location on chromosome are indicated). The list of species abbreviations: Aet — *Aegilops tauschii*, At — *Arabidopsis thaliana*, Bd — *Brachypodium distachyon*, Da — *Deschampsia antarctica*, Hv — *Hordeum vulgare*, Lr — *Lolium rigidum*, Ob — *Oryza brachyantha*, Os — *Oryza sativa*, Sb — *Sorghum bicolor*, Si — *Setaria italica*, Tae — *Triticum aestivum*, Zm — *Zea mays*. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Bootstrap values are shown next to the branches. The numbers of chromosomes (chr), where the corresponding genes are located, are indicated right to the accession numbers of the proteins. Brackets indicate grass ICE1 and ICE2 clusters.

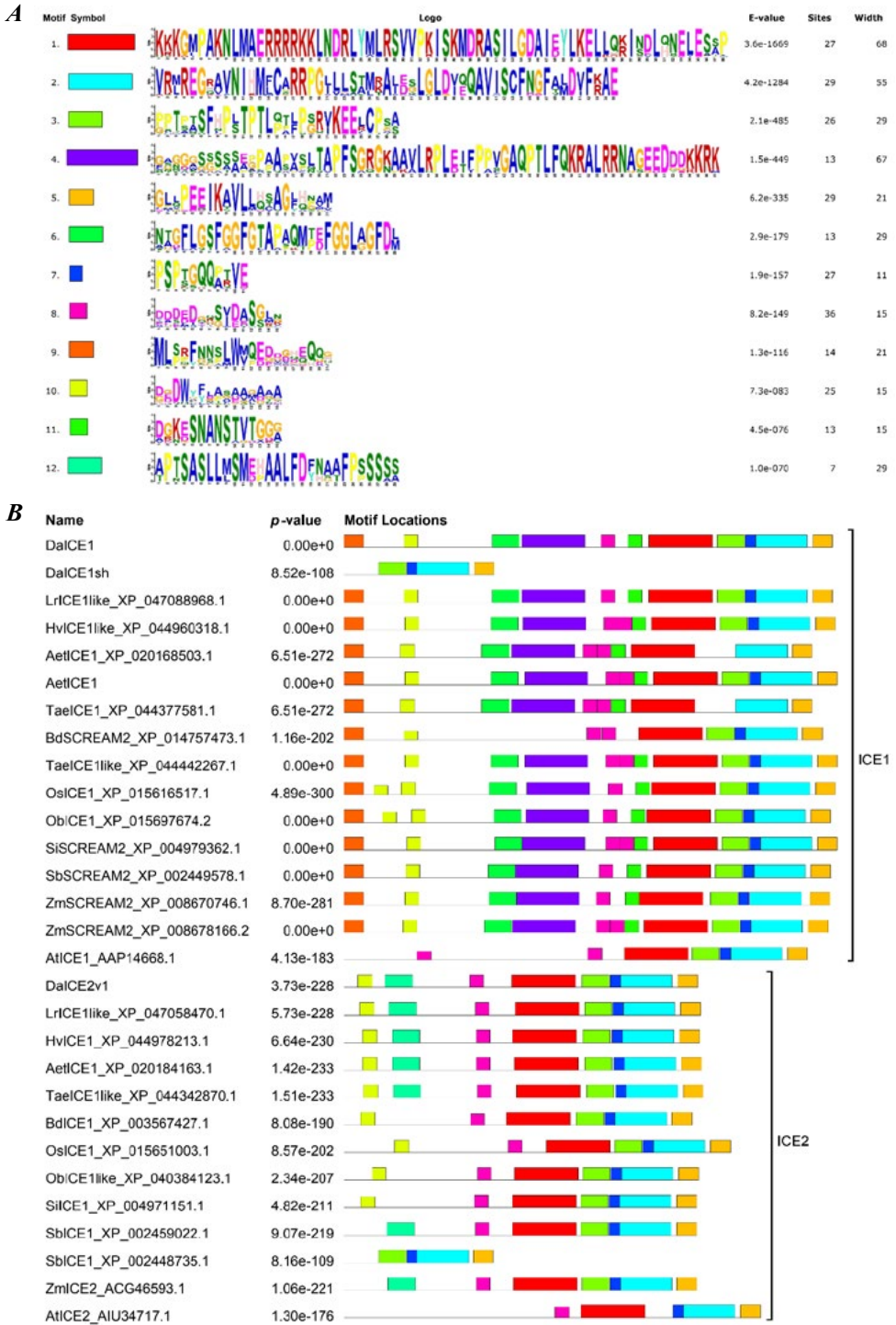

Fig. 4. Conserved motifs in ICE proteins (full information on the proteins is provided in the caption to Fig. 2): *A* — discovered motifs, *B* — motif locations.

the bHLH domain sequence as well as the ICE-specific sequence KMDRASILGDAIE/DYLKELL [48] was found in 27 proteins with the exception of truncated proteins DaICE1sh and SbICE1 XP_002448735.1. Motif 2 that contained the sequence specific for ACT-like domain was present in all studied proteins. Motif 4 that contained the ICE1-specific sequence GAQPTLFQKRA [12] was characteristic of the ICE1 group only.

The C-terminus of the proteins was similar in both groups in terms of the location and the occurrence of most of the motifs. Almost all proteins contained single motifs 1, 3, 7, 2, and 5 in this region. Motifs 3 and 7 were absent only in *Ae. tauschii* and *T. aestivum*. Additionally, one or two copies of motif 8 were present in all full-length proteins. Two short proteins, DaICE1sh and SbICE1 XP_002448735.1, contained only motifs 3, 7, 2, and 5. Motif 11 was specific only to the ICE1 subgroup, excluding *A. thaliana* and *B. distachyon*.

The N-terminus of the studied proteins was more polymorphic. Motif 10 was present in the majority of proteins in this region, whereas the remaining motifs were specific to individual subgroups. Motif 9 was found in all ICE1 proteins of grasses, whereas motifs 4 and 6 were characteristic of the same subgroup except for *B. distachyon*. Motif 12 was unique to ICE2, but it was not found in all proteins of this group.

The sites of post-translational modification of *D. antarctica* ICE TFs were predicted using the MusiteDeep program [38] (Table 2). DaICE1 had the highest number of glycosylation sites, 21 sites in the full-length and 14 sites in the short form, whereas in DaICE2 only 4 sites were found. The second most frequent were phosphorylation sites: 11, 5 and 9 sites were found in DaICE1, DaICE1sh and DaICE2, respectively. Six ubiquitination sites were present in DaICE1 and two in DaICE1sh and DaICE2 each. Also, two sumoylation sites were found in all proteins. Additionally, individual acetylation, methylation and hydroxylation sites were found in DaICE1 and DaICE2. Palmitoylation and pyrrolidone carboxylic acid binding sites were not detected (Table 2). A significant number of post-translational modification sites found is in good agreement with the prediction of the putative DaICE proteins as unstable in the absence of a stabilization mechanism (Table 1) using the ProtParam tool.

Our results are consistent with the published data on the key role of post-translational modifications of ICE in the regulation of cold stress response. This may be supported by the fact that in rice, cold stress resulted in increased levels of OsICE1 and OsICE2 proteins, but did not enhance the expression of their genes [48].

Table 2. **The number of predicted post-translational modification sites in putative *D. antarctica* ICE TFs**

	DaICE1	DaICE1sh	DaICE2v1,2
Phosphorylation	11	5	9
Glycosylation	21	14	4
Ubiquitination	6	2	2
SUMOylation	2	2	2
Acetylation	1	0	2
Methylation	1	0	2
Pyrrolidone carboxylic acid	0	0	0
Palmitoylation	0	0	0
Hydroxylation	1	0	3
In total	43	23	24

Previous studies have reported the involvement of sumoylation [24], ubiquitination [25] and phosphorylation [26, 27] in the regulation of these TFs. N-glycosylation of proteins is important for regulating their physicochemical properties, as it can affect protein conformation, stability, and biological activity [61].

A search for *cis*-elements in the promoter sequences of *DaICE1*/*DaICE1sh* and two variants of *DaICE2* (Table 3) was performed using the PlantCare online tools [43]. The multiple core promoter elements such as TATA-boxes and CAAT-boxes were identified, which are usually located in the functional position approximately 25–30 and 70 bp upstream of the starting point of transcription, respectively. TATA-boxes were found only in the proximal region of *DaICE1*/*DaICE1sh*, whereas in both variants of *DaICE2*, they were also present in the distal and basal regions. CAAT-boxes were located along the entire length of the promoters of all studied genes.

Numerous *cis*-elements involved in the response to abiotic and biotic stresses were found, constituting the largest group. All the promoters contained MYB and MYC TF binding sites, LTR (low temperature response), STRE (response to various stress factors including heat, osmotic stress, and shortage of nutrients), ABRE (abscisic acid response), ARE and GC-motif (anoxic-specific inducibility), WRE3 and WUN-motif (wound response). Only the promoter of *DaICE1*/*DaICE1sh* contained a W-box, which is a pathogen-inducible *cis*-element, a binding site for the WRKY TF and a gibberellin response element. In the promoter of the first *DaICE2* variant, a P-box associated with gibberellin response was found. The promoter of the second variant of the gene lacked the elements associated with the response to gibberellin. Instead, unlike the first variant promoter, it had a CGTCA motif and a TGACG motif involved in response to methyl jasmonate (MeJA). These two *cis*-elements were also present in *DaICE1*/*DaICE1sh*. Such differences in the structure of the two promoter variants can provide greater flexibility in the hormonal regulation of the *DaICE2* gene expression [47]. The abundance of *cis*-elements from response to stress group in the investigated promoters is well consistent with the function of *ICE* genes, specifically the response to abiotic stresses. In particular, the literature data show an increase in the expression of *RmICE1* in response not only to cold exposure, but also to dehydration and salinity [62]. The same was observed in *ZjICE1*, and additionally, a similar effect of abscisic acid treatment was found [59].

In all of the promoters, *cis*-elements associated with the response to light were found, including A-box, ACE, Box II, G-box, GATA-motif, I-box, Sp1, and TCT-motif. The interaction between light response and abiotic stress pathways has been demonstrated in higher plants [10]. The G-box was the most commonly identified *cis*-element from this group (Table 3). It has also been shown to be involved in a range of abiotic and biotic responses [63]. The *ICE* genes expression was reported to be induced also by light. For example, in *A. thaliana*, the expression level of the *ICE2* gene was approximately fourfold higher in light-grown compared to dark-grown plants, and the same was observed for the plants transferred from dark to light growth conditions. In contrast, the *ICE1* gene only showed a small increase in expression [23].

A Circadian *cis*-element was found in both *DaICE2* promoter variants but not in *DaICE1*/*DaICE1sh*. This is consistent with the data from Kurbidaeva *et al.* [23], where *ICE1* expression did not change during the day, whereas *ICE2* expression showed small circadian fluctuations.

Another group of *cis*-elements that we identified is 'Growth and development-related'. Most of the elements from this group, specifically GRA (involved in transcription in leaves), MSA-like (*cis*-element associated with cell cycle regulation), re2f-1 (E2F TF binding site, regulation of cell cycle), and TGA-element (auxin response), were found only in *DaICE1*/*DaICE1sh* promoter. A CAT-box associated with meristem-specific expression was found in all promoters, which may indicate the involvement of both DaICE1 and DaICE2 TFs in the protection of meristematic tissues. According to Kurbidaeva *et al.* [23], this function is specific only for ICE2 in *A. thaliana*, which the authors believe is related to the higher level of its expression in meristems and the presence of relevant *cis*-elements in the promoter.

The above-mentioned P-box and W-box can also be included in the latter group, since gibberellins regulate not only the adaptation of plants to biotic and abiotic stresses, but also play a critical role in the plant development [64].

Despite the differences in the promoter sequences of the three putative *DaICE* genes, all of them contain *cis*-elements, most of which are identical or similar in function. This is consistent with the similarity of the main functions of these TFs, such as regulation of cold resistance, stomatal development, and flowering time [23].

Building a three-dimensional structure of a protein can provide information on its function, binding sites, and can help to better characterize the proteins being studied. The first stage of prediction using Phyre2 web portal involves searching for homologous proteins in the PDB (Protein Data Bank) database followed by building a 3D model of homologous regions based on templates. The predicted ICE proteins of *D. antarctica* had the highest sequence homology with TF MYC2 of *A. thaliana* (PDB ID — 5GNJ) (homology confidence of 97.8-99.6 %) [65].

The percentage of amino acid residues modelled with >90 % confidence was 29 % for DaICE1, 69 % for DaICE1sh, and 37 % for DaICE2. These regions corresponded to the bHLH- and ACT-like domains. The remaining portion of the protein was predicted to have disordered structure. The percentage of disorder was high for DaICE1 and DaICE2, with values of 72 % and 64 %, respectively. Noteworthy, the disordered regions of proteins cannot be reliably predicted [41].

DaICE1 and DaICE2 proteins contain the bHLH domain (basic/helix-loop-helix) (Fig. 5A, C). As the name suggests, the basic region is located at the N-terminus of the domain and responsible for DNA binding to the *cis*-element E-box (CANNTG). In turn, the HLH region is involved in the formation of homodimers or heterodimers. It comprises two amphipathic α-helices separated by a loop that varies in both length and primary sequence [66]. The bHLH domains of *D. antarctica* ICE proteins were also predicted using the JPred4 server to encompass two α-helices [40].

The ACT-like domain that functions as a dimerization module [67, 68] is located at the

Table 3. **The types and number of *cis*-elements found in the promoter of *DaICE1* and *DaICE1sh* and two variants of *DaICE2* promoters**

Function	*Cis*-element, sequences	DaICE1, DaICEsh	DaICE2v1	DaICE2v2
Core promoter elements	TATA-box, TATA, ATTATA, TATATA, TATATAA, ATATAT, TATAA	2	7	7
	CAAT-box, CAAT, CCAAT, CAAAT	11	8	13
Response to abiotic and biotic stresses				
MYB TF binding site	MYB, CCGTTG, CAACCA, TAACTG	1	3	2
MYC TF binding site	MYC, CATTTG	2	1	1
Low temperature response	LTR, CCGAAA	1	1	2
Response to heat, osmotic stress, and nutrient shortage	STRE, AGGGG	3	2	1
Anoxic-specific inducibility	GC-motif, CCCCCG	1		
	ARE, AAACCA	2	2	1
Wound response	WRE3, CCACCT	1	2	1
	WUN-motif, AAATTACT			1
Abscisic acid responsiveness	ABRE, GACACGTGGC, TACGGTC, ACGTG, GCCGCGTGGC, CGTACGTGCA, GACACGTACGT, TACGTG, CACGTA	5	6	5
MeJA-responsiveness	CGTCA-motif, CGTCA	2		1
	TGACG-motif (as-1), TGACG	2		1
Gibberellin-responsive	P-box, CCTTTTG		1	
Pathogen-inducible, binding site for the WRKY TF, gibberellin response	W-box, TTGACC	1		
Response to light				
Elicitor- or light-mediated activation	A-box, CCGTCC	1		
Response to light	ACE, CTAACGTATT		1	
	Box II, TGGTAATAA		1	
	G-box CACGTT, CACGAC, CACGTC, TACGTG	3	3	3
	GATA-motif, AAGGATAAGG		1	1
	I-box, cGATAAGGCG	1		
	Sp1, GGGCGG	2		
	TCT-motif, TCTTAC		1	
Circadian	Circadian, CAAAGATATC		1	1
Growth and development-related				
Expression in the meristem	CAT-box, GCCACT	1	1	1

End of the Table 3

Function	*Cis*-element, sequences	DaICE1, DaICEsh	DaICE2v1	DaICE2v2
Transcription in leaves	GRA, CACTGGCCGCCC	1		
Cell cycle regulation	MSA-like, (T/C)C(T/C)AACGG(T/C)(T/C)A	1		
E2F TF binding site, cell cycle regulation	re2f-1, GCGGGAAA	2		
Auxin-responsive	TGA-element, AACGAC	1		
Total amount of types		22	17	16

C-terminus of DaICE proteins (Fig. 5). While its prototype bacterial ACT domain has a characteristic arrangement of four β-strands and two α-helices in a βαββαβ fold [69], the ACT-like domain of ICE forms a βαββα structure due to its shorter length [49]. The ACT-like domains of *D. antarctica* ICE proteins were predicted using the JPred4 server to also have the βαββα topology [40].

The expression of *DaICE* genes was measured using mRNA-Seq data sets from different organs, including roots, flowers, crowns, and leaves of plants, which were grown under laboratory conditions and taken from natural environment. Constitutive expression of all or most of the *DaICE* gene variants was found in all of the organs of plants grown in both laboratory and natural conditions (Fig. 6). However, the expression of the *DaICE2v2* variant was completely absent in the roots of both types of plants, and the expression of *DaICE1* was absent in the leaves of plants grown under laboratory conditions. *DaICE1* and *DaICE1sh* had the highest expression level in flowers, and *DaICE1* had the highest expression level in crowns of plants from natural conditions (Fig. 6).

The highest expression stability among the *DaICE* genes was demonstrated by both *DaICE2* variants. According to BestKeeper

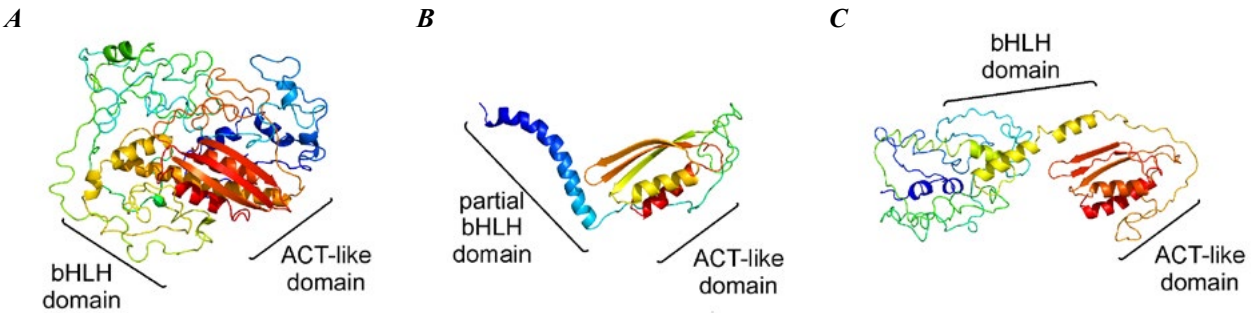

Fig. 5. Three-dimensional structures of DaICE proteins: *A* — DaICE1, *B* — DaICE1sh, *C* — DaICE2. The image of the amino acid chain is colored by rainbow from the N-terminus to the C-terminus.

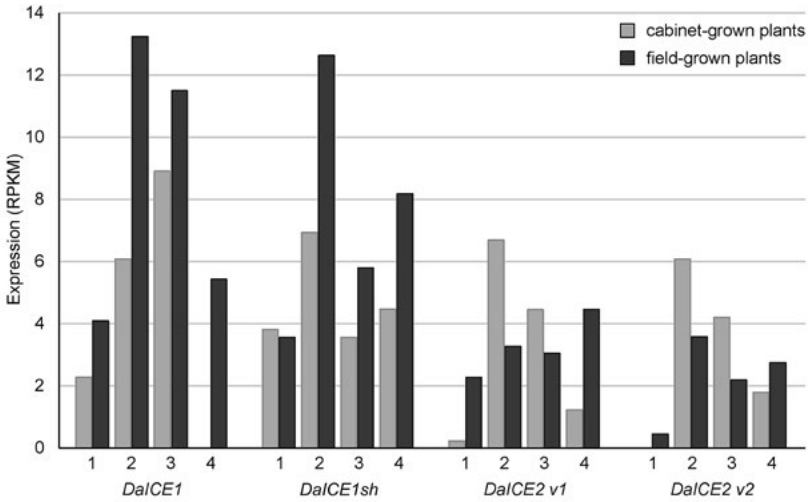

Fig. 6. Expression of *DaICE* genes in RPKM (reads per kilobase of transcript per million mapped reads) in various tissues of plants grown in two different environments (field and growth cabinet): 1 — roots, 2 — flowers, 3 — crowns, 4 — leaves.

data [45], the stability of gene expression for DaICE2v1 and DaICE2v2 was 1.41 and 1.56, respectively. The average stability of expression was determined for *DaICE1sh* (2.38), the lowest for *DaICE1* (3.41). For all of the genes, the values exceeded one, thus indicating that their expression in different tissues under different growth conditions was unstable [45].

Our results are consistent with previous studies showing constitutive expression of *ICE1* and *ICE2* genes in various species [9, 12, 56]. In the plants grown in natural extreme environment, the expression levels of all *DaICE1* variants were higher compared to the plants grown under laboratory conditions. The *DaICE2* variants did not show a clear relationship with growth conditions; we found increased expression in the roots and leaves from wild-grown plants, and in the flowers and crowns from plants grown in growth cabinet (Fig. 6). Several studies have reported an increase in the expression level of *ICE1* under cold stress [9, 47, 55, 56, 58, 62, 70, 71]. However, there are also reports on maintaining a stable level of *ICE1* expression [12, 47, 48]. For the *ICE2* gene, other researchers have also observed no significant change in expression level under cold stress [12, 48, 55]. The detection of *ICE* gene transcripts in plants grown at normal temperature and under cold exposure suggests that their activity is regulated at the protein level by post-translational modification.

The expression of *DaICE* genes was found to be the lowest in the roots of plants grown in both environments and in the leaves of wild-grown plants, while it was relatively higher in flowers, crowns, and leaves of wild-grown plants (Fig. 6). Literature data confirm that *ICE1* expression may vary in different tissues, but the observed patterns differ. For example, higher expression was observed in roots and shoots compared to leaves of *Raphanus sativus* [71], and in leaves and shoots compared to roots and flowers of *A. thaliana* [9] and *Hevea brasiliensis* [56].

Conclusions

ICE TFs play a key role in the plant tolerance to low temperatures. They are the initial link in the main cold stress signaling pathway in plants, i.e. the ICE–CBF–COR pathway, and also integrate cold-induced signals into physiological processes such as stomatal development and flowering. Currently, little is known about the TFs involved in cold stress response in the extremophile grass *D. antarctica*. We predicted *in silico* the *DaICE1* and *DaICE2* genes for the first time and demonstrated the presence of characteristic bHLH and ACT-like domains and conserved motifs in the coded proteins as well as their homology to the *ICE* orthologous genes of other grass species.

We found two variants of the *DaICE1* gene: a normal one that encodes a full-length protein, and a variant with a large deletion. The protein encoded by the latter gene lost the N-terminus and part of the bHLH domain, but retained an ACT-like domain involved in the formation of heterodimers with other TFs from the bHLH group during the switch from proliferation to differentiation in stomatal development. We also found a similar gene in sorghum. Such truncated proteins can potentially act on the mechanism of microProteins, in particular in the negative regulation of stomatal differentiation processes.

We characterized the putative DaICE TFs using the bioinformatics methods. We predicted their nuclear localization, physicochemical properties, three-dimensional structure, and potential sites for post-translational modifications. We also studied the organization of promoters and found that the *cis*-elements involved in the response to abiotic stress formed the largest group.

We detected gene expression in different tissues and demonstrated that the expression of the *DaICE1* variants was increased under natural extreme conditions compared to the laboratory conditions. These findings suggest that these TFs may serve as regulators of the *CBF* gene expression during cold stress. Our data contribute to the study of the cold stress response in grasses and provide a theoretical basis for the future efforts to increase the stress resistance of agricultural crops.

Funding

The study was performed within the framework of research project “Genetic, physiological, and biochemical mechanisms for plant adaptation to extreme environmental conditions” for 2021-2025, No. 0120U105249.

REFERENCES

1. *Parnikoza I, Kozeretska I, Kunakh V.* Vascular Plants of the Maritime Antarctic: Origin and Adaptation. *Am J Plant Sci.* 2011; **2**(3):381–95.
2. *Byun MY, Lee J, Cui LH, Kang Y, Oh TK, Park H, Lee H, Kim WT.* Constitutive expression of *DaCBF7*, an Antarctic vascular plant *Deschampsia antarctica* CBF homolog, resulted in improved cold tolerance in transgenic rice plants. *Plant Sci.* 2015; **236**:61–74.
3. *Byun MY, Cui LH, Lee J, Park H, Lee A, Kim WT, Lee H.* Identification of Rice Genes Associated With Enhanced Cold Tolerance by Comparative Transcriptome Analysis With Two Transgenic Rice Plants Overexpressing *DaCBF4* or *DaCBF7*, Isolated From Antarctic Flowering Plant *Deschampsia antarctica*. *Front Plant Sci.* 2018; **9**:601.
4. *Byun MY, Cui LH, Lee A, Oh HG, Yoo YH, Lee J, Kim WT, Lee H.* Abiotic Stress-Induced Actin-Depolymerizing Factor 3 From *Deschampsia antarctica* Enhanced Cold Tolerance When Constitutively Expressed in Rice. *Front Plant Sci.* 2021; **12**:734500.

5. *Lee J, Noh EK, Choi HS, Shin SC, Park H, Lee H.* Transcriptome sequencing of the Antarctic vascular plant *Deschampsia antarctica* Desv. *under abiotic stress. Planta.* 2013; **237**(3):823–36.
6. *Lee J, Kang Y, Shin SC, Park H, Lee H.* Combined analysis of the chloroplast genome and transcriptome of the Antarctic vascular plant *Deschampsia antarctica* Desv. *PLoS One.* 2014; **9**(3):e92501.
7. *Bublyk OM, Andreev IO, Kunakh VA.* Bioinformatic prediction of genes of cold-induced transcription factors CBF/DREB1 and DREB4 in *Deschampsia antarctica* Desv. *UAJ.* 2016; **15**:81–95.
8. *Thomashow MF.* Molecular basis of plant cold acclimation: insights gained from studying the CBF cold response pathway. *Plant Physiol.* 2010; **154**(2):571–7.
9. *Chinnusamy V, Ohta M, Kanrar S, Lee BH, Hong X, Agarwal M, Zhu JK.* ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *Arabidopsis*. *Genes Dev.* 2003; **17**(8):1043–54.
10. *Wang DZ, Jin YN, Ding XH, Wang WJ, Zhai SS, Bai LP, Guo ZF.* Gene Regulation and Signal Transduction in the ICE-CBF-COR Signaling Pathway during Cold Stress in Plants. *Biochemistry (Mosc).* 2017; **82**(10):1103–17.
11. *Fursova OV, Pogorelko GV, Tarasov VA.* Identification of *ICE2*, a gene involved in cold acclimation which determines freezing tolerance in *Arabidopsis thaliana*. *Gene.* 2009; **429**(1–2):98–103.
12. *Badawi M, Reddy YV, Agharbaoui Z, Tominaga Y, Danyluk J, Sarhan F, Houde M.* Structure and functional analysis of wheat *ICE* (inducer of CBF expression) genes. *Plant Cell Physiol.* 2008; **49**(8):1237–49.
13. *Feng XM, Zhao Q, Zhao LL, Qiao Y, Xie XB, Li HF, Yao YX, You CX, Hao YJ.* The cold-induced basic helix-loop-helix transcription factor gene *MdCIbHLH1* encodes an ICE-like protein in apple. *BMC Plant Biol.* 2012; **12**:22.
14. *Verma RK, Kumar VVS, Yadav SK, Kumar TS, Rao MV, Chinnusamy V.* Overexpression of *Arabidopsis ICE1* enhances yield and multiple abiotic stress tolerance in indica rice. *Plant Signal Behav.* 2020; **15**(11):1814547.
15. *Feng HL, Ma NN, Meng X, Zhang S, Wang JR, Chai S, Meng QW.* A novel tomato MYC-type ICE1-like transcription factor, SlICE1a, confers cold, osmotic and salt tolerance in transgenic tobacco. *Plant Physiol Biochem.* 2013; **73**:309–20.
16. *Li J, Wang L, Zhu W, Wang N, Xin H, Li Sh.* Characterization of two *VvICE1* genes isolated from 'Muscat Hamburg' grapevine and their effect on the tolerance to abiotic stresses. *Sci Hortic.* 2014; **165**:266–73. https://doi.org/10.1016/j.scienta.2013.11.002.
17. *Kanaoka MM, Pillitteri LJ, Fujii H, Yoshida Y, Bogenschutz NL, Takabayashi J, Zhu JK, Torii KU.* SCREAM/ICE1 and SCREAM2 specify three cell-state transitional steps leading to *Arabidopsis* stomatal differentiation. *Plant Cell.* 2008; **20**(7):1775–85.
18. *Lee JH, Jung JH, Park CM.* INDUCER OF CBF EXPRESSION 1 integrates cold signals into FLOWERING LOCUS C-mediated flowering pathways in *Arabidopsis*. *Plant J.* 2015; **84**(1):29–40.
19. *MacGregor DR, Zhang N, Iwasaki M, Chen M, Dave A, Lopez-Molina L, Penfield S.* ICE1 and ZOU determine the depth of primary seed dormancy in *Arabidopsis* independently of their role in endosperm development. *Plant J.* 2019; **98**(2):277–90.
20. *Hu Y, Han X, Yang M, Zhang M, Pan J, Yu D.* The Transcription Factor INDUCER OF CBF EXPRESSION1 Interacts with ABSCISIC ACID INSENSITIVE5 and DELLA Proteins to Fine-Tune Abscisic Acid Signaling during Seed Germination in *Arabidopsis*. *Plant Cell.* 2019; **31**(7):1520–38.
21. *Denay G, Creff A, Moussu S, Wagnon P, Thévenin J, Gérentes MF, Chambrier P, Dubreucq B, Ingram G.* Endosperm breakdown in *Arabidopsis* requires heterodimers of the basic helix-loop-helix proteins ZHOUPI and INDUCER OF CBP EXPRESSION 1. *Development.* 2014; **141**(6):1222–7.
22. *Wei D, Liu M, Chen H, Zheng Y, Liu Y, Wang X, Yang S, Zhou M, Lin J.* INDUCER OF CBF EXPRESSION 1 is a male fertility regulator impacting anther dehydration in *Arabidopsis*. *PLoS Genet.* 2018; **14**(10):e1007695.
23. *Kurbidaeva A, Ezhova T, Novokreshchenova M. Arabidopsis thaliana ICE2* gene: phylogeny, struc-

tural evolution and functional diversification from *ICE1*. *Plant Sci.* 2014; **229**:10–22.
24. *Miura K, Jin JB, Lee J, Yoo CY, Stirm V, Miura T, Ashworth EN, Bressan RA, Yun DJ, Hasegawa PM.* SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A expression and freezing tolerance in *Arabidopsis*. *Plant Cell.* 2007; **19**(4):1403–14.
25. *Dong CH, Agarwal M, Zhang Y, Xie Q, Zhu JK.* 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 U S A.* 2006; **103**(21):8281–6.
26. *Ding Y, Li H, Zhang X, Xie Q, Gong Z, Yang S.* OST1 kinase modulates freezing tolerance by enhancing ICE1 stability in *Arabidopsis*. *Dev Cell.* 2015; **32**(3):278–89.
27. *Zhang Z, Li J, Li F, Liu H, Yang W, Chong K, Xu Y.* OsMAPK3 Phosphorylates OsbHLH002/OsICE1 and Inhibits Its Ubiquitination to Activate OsTPP1 and Enhances Rice Chilling Tolerance. *Dev Cell.* 2017; **43**(6):731–43.
28. *Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL.* BLAST+: architecture and applications. *BMC Bioinformatics.* 2009; **10**:421.
29. *Kumar S, Stecher G, Li M, Knyaz C, Tamura K.* MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol.* 2018; **35**(6):1547–9.
30. *Hu B, Jin J, Guo AY, Zhang H, Luo J, Gao G.* GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics.* 2015; **31**(8):1296–7.
31. *Edgar RC.* MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; **32**(5):1792–7.
32. *Jones DT, Taylor WR, Thornton JM.* The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci.* 1992; **8**(3):275–82.
33. *Nicholas KB, Nicholas HBJ, Deerfield DW.* GeneDoc: analysis and visualization of genetic variation. *EMBNEW News.* 1997; **4**:14–6.
34. *Geer LY, Domrachev M, Lipman DJ, Bryant SH.* CDART: protein homology by domain architecture. *Genome Res.* 2002; **12**(10):1619–23.
35. *Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A.* Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) The proteomics protocols handbook. Louisville: Humana, 2005. — 571–607 pp.
36. *Chou KC, Shen HB.* Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS One.* 2010; **5**(6):e11335.
37. *Guo Y, Yang Y, Huang Y, Shen HB.* Discovering nuclear targeting signal sequence through protein language learning and multivariate analysis. *Anal Biochem.* 2020; **591**:113565.
38. *Wang D, Liu D, Yuchi J, He F, Jiang Y, Cai S, Li J, Xu D.* MusiteDeep: a deep-learning based webserver for protein post-translational modification site prediction and visualization. *Nucleic Acids Res.* 2020; **48**(W1):W140–W146.
39. *Bailey TL, Johnson J, Grant CE, Noble WS.* The MEME Suite. *Nucleic Acids Res.* 2015; **43**(W1):W39–49.
40. *Drozdetskiy A, Cole C, Procter J, Barton GJ.* JPred4: a protein secondary structure prediction server. *Nucleic Acids Res.* 2015; **43**(W1):W389–94.
41. *Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ.* The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc.* 2015; **10**(6):845–58.
42. *Ittisoponpisan S, Islam SA, Khanna T, Alhuzimi E, David A, Sternberg MJE.* Can Predicted Protein 3D Structures Provide Reliable Insights into whether Missense Variants Are Disease Associated? *J Mol Biol.* 2019; **431**(11):2197–212.
43. *Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S.* PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 2002; **30**(1):325–7.
44. *Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B.* Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods.* 2008; **5**(7):621–8.
45. *Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP.* Determination of stable housekeeping genes, dif-

ferentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnol Lett.* 2004; **26**(6):509–15.
46. *Rahman MA, Moody MA, Nassuth A.* Grape contains 4 *ICE* genes whose expression includes alternative polyadenylation, leading to transcripts encoding at least 7 different ICE proteins. *Environ Exp Bot.* 2014; **106**:70–8.
47. *Kashyap P, Deswal R.* Two ICE isoforms showing differential transcriptional regulation by cold and hormones participate in *Brassica juncea* cold stress signaling. *Gene.* 2019; **695**:32–41.
48. *Nakamura J, Yuasa T, Huong TT, Harano K, Tanaka S, Iwata T, Phan T, Iwaya M.* Rice homologs of inducer of *CBF* expression (OsICE) are involved in cold acclimation. *Plant Biotechnol.* 2011; **28**(3):303–9.
49. *Seo H, Sepuru KM, Putarjunan A, Aguirre L, Burrows BA, Torii KU.* Intragenic suppressors unravel the role of the SCREAM ACT-like domain for bHLH partner selectivity in stomatal development. *Proc Natl Acad Sci U S A.* 2022; **119**(9):e2117774119.
50. *Staudt AC, Wenkel S.* Regulation of protein function by 'microProteins'. *EMBO Rep.* 2011; **12**(1):35–42.
51. *Seo PJ, Hong SY, Kim SG, Park CM.* Competitive inhibition of transcription factors by small interfering peptides. *Trends Plant Sci.* 2011; **16**(10):541–9.
52. *Magnani E, de Klein N, Nam HI, Kim JG, Pham K, Fiume E, Mudgett MB, Rhee SY.* A comprehensive analysis of microProteins reveals their potentially widespread mechanism of transcriptional regulation. *Plant Physiol.* 2014; **165**(1):149–59.
53. *Xu Z, Zhou G.* Responses of leaf stomatal density to water status and its relationship with photosynthesis in a grass. *J Exp Bot.* 2008; **59**(12):3317–25.
54. *de Oliveira JPV, Duarte VP, de Castro EM, Magalhães PC, Pereira FJ.* Stomatal cavity modulates the gas exchange of *Sorghum bicolor* (L.*) Moench.* grown under different water levels. *Protoplasma.* 2022; **259**(4):1081–97.
55. *Xu W, Jiao Y, Li R, Zhang N, Xiao D, Ding X, Wang Z.* Chinese wild-growing *Vitis amurensis ICE1* and *ICE2* encode MYC-type bHLH transcription activators that regulate cold tolerance in *Arabidopsis*. *PLoS One.* 2014; **9**(7):e102303.
56. *Deng XM, Wang, JX, Li Y, Wang J, Tian W-M.* Characterization of a cold responsive *HbICE1* gene from rubber trees. *Trees.* 2017; **31**(1):137–47.
57. *Lu X, Yang L, Yu M, Lai J, Wang C, McNeil D, Zhou M, Yang C.* A novel *Zea mays* ssp. *mexicana L.* MYC-type ICE-like transcription factor gene *ZmmICE1*, enhances freezing tolerance in transgenic *Arabidopsis thaliana*. *Plant Physiol Biochem.* 2017; **113**:78–88.
58. *Jin Y, Zhai S, Wang W, Ding X, Guo Z, Bai L, Wang S.* Identification of genes from the ICE-CBF-COR pathway under cold stress in *Aegilops-Triticum* composite group and the evolution analysis with those from Triticeae. *Physiol Mol Biol Plants.* 2018; **24**(2):211–29.
59. *Zuo ZF, Kang HG, Park MY, Jeong H, Sun HJ, Song PS, Lee HY. Zoysia japonica* MYC type transcription factor ZjICE1 regulates cold tolerance in transgenic *Arabidopsis*. *Plant Sci.* 2019; **289**:110254.
60. *Nakano T, Suzuki K, Fujimura T, Shinshi H.* Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. *Plant Physiol.* 2006; **140**(2):411–32.
61. *Nagashima Y, von Schaewen A, Koiwa H.* Function of N-glycosylation in plants. *Plant Sci.* 2018; **274**:70–9.
62. *Luo P, Li Z, Chen W, Xing W, Yang J, Cui Y.* Overexpression of *RmICE1*, a bHLH transcription factor from *Rosa multiflora*, enhances cold tolerance via modulating ROS levels and activating the expression of stress-responsive genes. *Environ Exp Bot.* 2020; **178**:104160.
63. *Zhang N, McHale LK, Finer JJ.* Changes to the core and flanking sequences of G-box elements lead to increases and decreases in gene expression in both native and synthetic soybean promoters. *Plant Biotechnol J.* 2019; **17**(4):724–35.
64. *Gao S, Chu C.* Gibberellin Metabolism and Signaling: Targets for Improving Agronomic Performance of Crops. *Plant Cell Physiol.* 2020; **61**(11):1902–11.
65. *Lian TF, Xu YP, Li LF, Su XD.* Crystal Structure of Tetrameric *Arabidopsis* MYC2 Reveals the Mechanism of Enhanced Interaction with DNA. *Cell Rep.* 2017; **19**(7):1334–42.
66. *Anthony-Cahill SJ, Benfield PA, Fairman R, Wasserman ZR, Brenner SL, Stafford WF 3rd, Alten-*

bach C, Hubbell WL, DeGrado WF. Molecular characterization of helix-loop-helix peptides. *Science.* 1992; **255**(5047):979–83.
67. *Feller A, Hernandez JM, Grotewold E.* An ACT-like domain participates in the dimerization of several plant basic-helix-loop-helix transcription factors. *J Biol Chem.* 2006; **281**(39):28964–74.
68. *Kong Q, Pattanaik S, Feller A, Werkman JR, Chai C, Wang Y, Grotewold E, Yuan L.* Regulatory switch enforced by basic helix-loop-helix and ACT-domain mediated dimerizations of the maize transcription factor R. *Proc Natl Acad Sci U S A.* 2012; **109**(30): E2091–7.
69. *Chipman DM, Shaanan B.* The ACT domain family. *Curr Opin Struct Biol.* 2001; **11**(6):694–700.
70. *Guo J, Ren Y, Tang Z, Shi W, Zhou M.* Characterization and expression profiling of the ICE-CBF-COR genes in wheat. *PeerJ.* 2019; **7**:e8190.
71. *Man L, Xiang D, Wang L, Zhang W, Wang X, Qi G.* Stress-responsive gene *RsICE1* from *Raphanus sativus* increases cold tolerance in rice. *Protoplasma.* 2017; **254**(2):945–56.

Передбачення та аналіз стрес-індукованих транскрипційних факторів ICE у *Deschampsia antarctica*

О. М. Бублик, І. О. Андрєєв, В. А. Кунах

Мета. Дослідження молекулярно-генетичних аспектів стресостійкості рослини-екстремофіла *D. antarctica* на прикладі аналізу транскрипційних факторів ICE, залучених до відповіді на холодовий стрес, регуляції розвитку продихів і часу цвітіння. **Методи.** Гени *DaICE* збирали із послідовностей, гомологічних до генів-ортологів, з бази даних SRA. Передбачені гени та білки аналізували біоінформатичними методами. Генну експресію аналізували з використанням даних сиквенування РНК. **Результати.** Знайдено два варіанти гена *DaICE1*: типовий, який кодує повнорозмірний білок, та варіант із великою делецією, що призвела до втрати N-кінця та частини bHLH-домена у білкового продукта. Для гена *DaICE2* знайдено два близькі варіанти з чотирма однонуклеотидними поліморфізмами, три з яких є синонімічними, в кодувальній ділянці та більш відмінними промоторними послідовностями. Досліджено організацію промоторів та інтрон-екзонну структуру генів. Для передбачених ТФ визначено фізико-хімічні властивості, показано локалізацію в ядрі, можливість посттрансляційних модифікацій, проведено філогенетичний аналіз, пошук консервативних мотивів, передбачено тривимірну структуру. Виявлено експресію генів у більшості або всіх тканинах та показано підвищену експресію варіантів *DaICE1* в рослинах з природних умов, порівняно з лабораторними. **Висновки.** *DaICE1* та *DaICE2* є кандидатами на роль регуляторів експресії генів CBF під час холодового стресу у *D. antarctica*.

Ключові слова: антарктичні рослини, транскрипційні фактори, передбачення *in silico*, абіотичний стрес, стійкість.

Received 27.04.2023