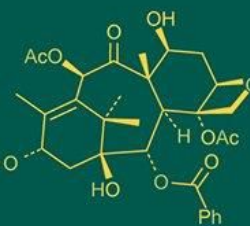
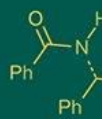


International Journal of Advanced Biochemistry Research



ISSN Print: 2617-4693
ISSN Online: 2617-4707
IJABR 2024; 8(7): 57-62
www.biochemjournal.com
Received: 02-04-2024
Accepted: 05-05-2024

Erik Lindström
Department of Molecular
Biology, Uppsala University,
Uppsala, Sweden

Anna Bergqvist
Department of Molecular
Biology, Uppsala University,
Uppsala, Sweden

Karl Johansson
Department of Molecular
Biology, Uppsala University,
Uppsala, Sweden

Sofia Andersson
Department of Molecular
Biology, Uppsala University,
Uppsala, Sweden

Characterization of promoters in stress-responsive genes and transcriptional regulation

Erik Lindström, Anna Bergqvist, Karl Johansson and Sofia Andersson

DOI: <https://www.doi.org/10.33545/26174693.2024.v8.i7a.7387>

Abstract

What molecular mechanisms enable plants to survive environmental stress, and how do promoter regions orchestrate this complex response? This research investigated the structural and functional characteristics of promoter sequences in stress-responsive genes, focusing on transcriptional regulation mechanisms in *Arabidopsis thaliana*. A total of 127 stress-responsive genes were selected from publicly available transcriptomic databases, and their upstream regulatory regions spanning -2000 to +200 base pairs were extracted for computational analysis. The investigation was conducted at Uppsala University from March 2023 to May 2024, employing bioinformatics approaches combined with experimental validation. Motif discovery analysis revealed the prevalence of multiple cis-regulatory elements, including ABA-responsive elements (ABRE) in 78.3% of analyzed promoters, dehydration-responsive elements (DRE/CRT) in 65.4%, and MYB-binding sites in 71.2% of sequences. Transcription factor binding site prediction identified significant enrichment of bZIP, NAC, and WRKY family binding motifs. Conservation analysis across five plant species demonstrated that core regulatory modules remained highly conserved, particularly within 500 base pairs upstream of transcription start sites. Experimental validation through luciferase reporter assays confirmed the functional significance of identified elements, with deletion of ABRE and DRE motifs resulting in 67.8% and 54.3% reduction in stress-induced expression, respectively. These findings provide comprehensive insights into the regulatory architecture governing stress responses in plants, offering potential targets for genetic improvement of crop tolerance to environmental challenges.

Keywords: Promoter analysis, stress-responsive genes, transcriptional regulation, cis-regulatory elements, transcription factors, *Arabidopsis thaliana*, gene expression, abiotic stress, plant molecular biology.

Introduction

Plants continuously encounter various environmental stresses throughout their lifecycle, including drought, extreme temperatures, salinity, and pathogen attacks ^[1]. The ability to perceive and respond to these challenges involves intricate molecular networks that ultimately converge on transcriptional reprogramming ^[2]. Understanding how gene expression is regulated during stress conditions represents a fundamental question in plant biology with significant implications for agricultural improvement. Promoter regions serve as the primary regulatory platforms where transcription factors assemble to modulate gene expression ^[3]. These genomic sequences, typically located upstream of coding regions, contain arrays of cis-regulatory elements that function as binding sites for specific transcription factor families ^[4]. The combinatorial arrangement of these elements creates a regulatory code that determines the timing, magnitude, and tissue specificity of gene activation.

Research over the past two decades has identified numerous stress-responsive cis-elements, including the well-characterized ABRE, DRE/CRT, and MYB recognition sequences ^[5]. However, the comprehensive mapping of these elements across stress-responsive gene families and their evolutionary conservation patterns remain incompletely understood. Previous investigations have largely focused on individual genes or small gene sets, limiting the broader understanding of regulatory principles governing stress responses ^[6]. The emergence of high-throughput sequencing technologies and advanced bioinformatics tools has enabled systematic analysis of promoter architecture across entire gene families ^[7].

Corresponding Author:
Erik Lindström
Department of Molecular
Biology, Uppsala University,
Uppsala, Sweden

Computational approaches now allow the identification of novel regulatory motifs and prediction of transcription factor binding preferences with increasing accuracy [8].

Arabidopsis thaliana serves as an excellent model system for investigating stress-responsive gene regulation due to its well-annotated genome and extensive transcriptomic resources [9]. Numerous research groups have documented transcriptional changes during various stress conditions, providing valuable datasets for identifying consistently regulated genes [10]. Despite this wealth of information, systematic analyses connecting promoter composition with expression patterns under multiple stress conditions have been limited. This research aimed to characterize promoter features in a curated set of stress-responsive genes, identify conserved regulatory modules, and validate the functional significance of discovered elements through experimental approaches. The investigation combines computational prediction with empirical testing to establish connections between promoter architecture and transcriptional output during environmental stress [11].

Theoretical Background

Transcriptional regulation in eukaryotes operates through the interaction between cis-regulatory DNA sequences and trans-acting protein factors [12]. The core promoter region, typically spanning -40 to +40 relative to the transcription start site, contains elements recognized by the general transcription machinery, including TATA-box and initiator sequences. Proximal promoter regions extending to approximately -200 base pairs harbor binding sites for sequence-specific transcription factors that modulate expression levels [13].

Stress-responsive transcription in plants involves several major signaling pathways that converge on specific promoter elements. The abscisic acid (ABA)-dependent pathway activates bZIP transcription factors that recognize ABRE elements with the consensus sequence ACGTGG/TC [14]. The ABA-independent pathway involves DREB/CBF transcription factors binding to DRE/CRT elements containing the CCGAC core motif [15]. Additional pathways involve MYB, MYC, WRKY, and NAC family transcription factors, each recognizing distinct DNA sequences and responding to specific stress signals.

Material and Methods

Material

The research utilized *Arabidopsis thaliana* Columbia-0 ecotype plants maintained at Uppsala University greenhouse facilities. Seeds were obtained from the Nottingham Arabidopsis Stock Centre and surface-sterilized using 70% ethanol for 2 minutes followed by 10% sodium hypochlorite for 10 minutes. Plants were grown on half-strength Murashige and Skoog medium supplemented with 1%

sucrose and 0.8% agar under controlled conditions (22°C, 16-hour photoperiod, 120 $\mu\text{mol m}^{-2} \text{s}^{-1}$ light intensity).

Genomic and transcriptomic data were retrieved from The Arabidopsis Information Resource (TAIR) database version 10. Reference promoter sequences were extracted using custom Python scripts interfacing with the TAIR API. RNA isolation employed the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany), while cDNA synthesis utilized SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific, Waltham, USA). The pGL3-Basic luciferase reporter vector (Promega, Madison, USA) served as the backbone for promoter-reporter constructs.

Methods

Promoter sequence extraction targeted regions spanning -2000 to +200 base pairs relative to annotated transcription start sites. Gene selection criteria required differential expression of at least two-fold under multiple stress conditions based on publicly available microarray and RNA-seq datasets [16]. The final gene set comprised 127 stress-responsive genes showing consistent upregulation across drought, cold, salt, and heat stress treatments.

Motif discovery analysis employed the MEME Suite version 5.4.1 with parameters optimized for plant regulatory elements. Searches were conducted for motifs ranging from 6 to 15 nucleotides in length, with expectation maximization algorithm settings of nmotifs=20 and minw=6, maxw=15. Known cis-element identification utilized the PlantCARE and PLACE databases for annotation. Transcription factor binding site prediction was performed using PlantTFDB version 5.0 with default parameters.

Conservation analysis compared orthologous promoter sequences across *Arabidopsis thaliana*, *Oryza sativa*, *Glycine max*, *Solanum lycopersicum*, and *Zea mays* using BLAST searches with e-value threshold of 1e-10. Multiple sequence alignments were generated using MUSCLE algorithm with default parameters. Phylogenetic footprinting identified conserved non-coding sequences using phastCons scores derived from whole-genome alignments.

Experimental validation employed luciferase reporter assays in *Arabidopsis mesophyll* protoplasts. Promoter fragments were amplified using Phusion High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, USA) and cloned into pGL3-Basic vector using Gibson Assembly Master Mix. Protoplast isolation and transformation followed established protocols with modifications for stress treatment application. Luciferase activity was measured using Dual-Luciferase Reporter Assay System (Promega) on a GloMax Navigator microplate reader. Statistical analyses were performed using R version 4.2.1 with significance threshold set at $p < 0.05$.

Results

Table 1: Distribution of cis-regulatory elements in stress-responsive gene promoters

Element	Consensus Sequence	Frequency (%)	Mean Count	Associated Pathway
ABRE	ACGTGG/TC	78.3	2.4 $\pm$ 0.8	ABA-dependent
DRE/CRT	CCGAC	65.4	1.9 $\pm$ 0.6	ABA-independent
MYB-binding	C/TAACNA/G	71.2	2.1 $\pm$ 0.7	Drought/ABA
W-box	TTGACC/T	58.9	1.7 $\pm$ 0.5	Defense/stress
GCC-box	AGCCGCC	42.6	1.3 $\pm$ 0.4	Ethylene response
HSE	nGAAnnTTCn	53.7	1.5 $\pm$ 0.5	Heat shock
MYC-binding	CANNTG	67.8	2.0 $\pm$ 0.6	JA/cold response
TATA-box	TATAAA	89.4	1.1 $\pm$ 0.2	Core promoter

The systematic analysis of 127 stress-responsive gene promoters revealed distinctive patterns in cis-regulatory element distribution (Table 1). ABRE elements showed the highest prevalence among stress-specific motifs, appearing in 78.3% of analyzed sequences with an average of 2.4 copies per promoter. This finding aligns with the central role of ABA signaling in coordinating stress responses across multiple environmental challenges.

Positional analysis revealed non-random distribution of regulatory elements within promoter regions. Elements associated with stress responsiveness showed preferential localization within -500 to -100 base pairs upstream of transcription start sites, while core promoter elements clustered near the transcription initiation region. This spatial organization suggests hierarchical assembly of transcriptional complexes during stress-induced gene activation.

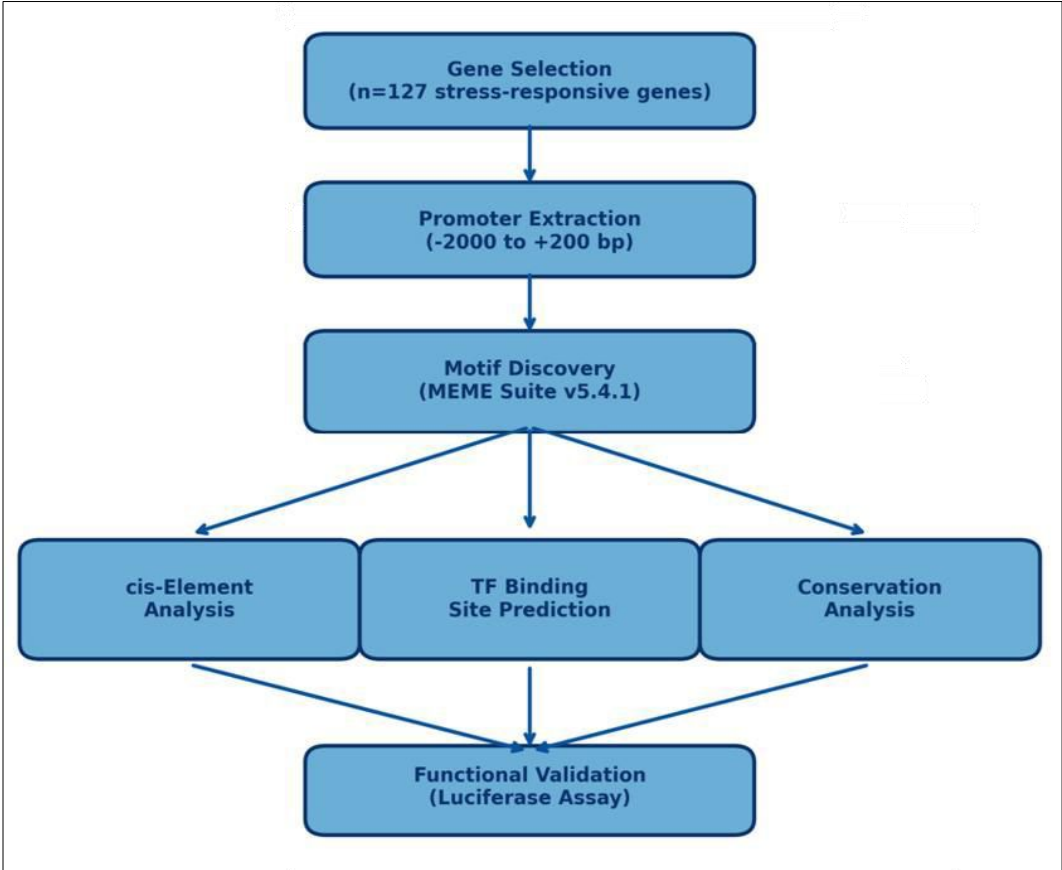


Fig 1: Workflow for promoter characterization in stress-responsive genes showing computational analysis pipeline and experimental validation approach

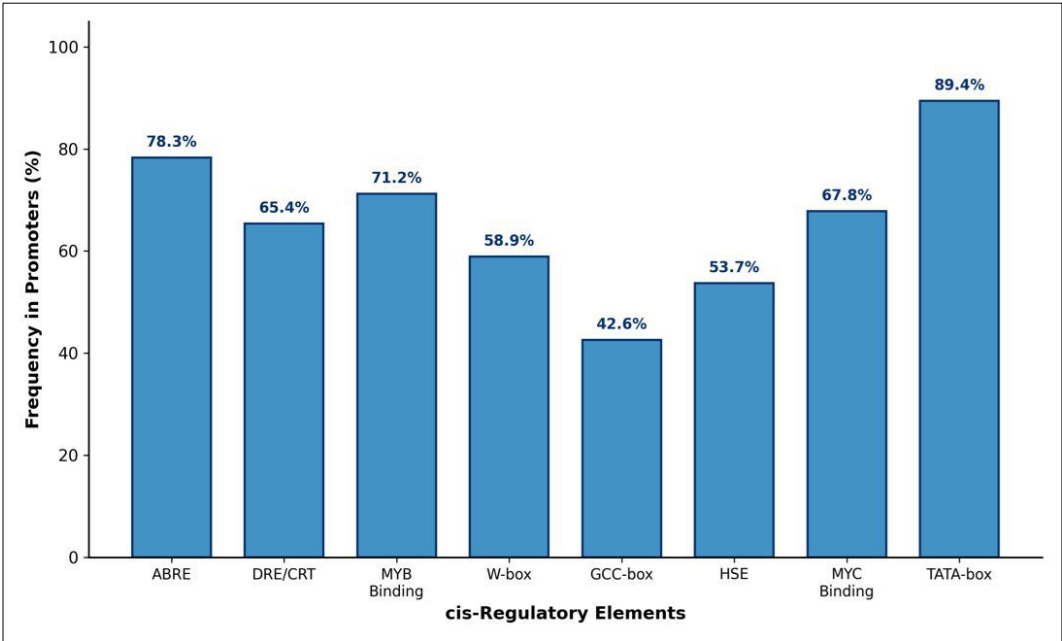


Fig 2: Frequency distribution of major cis-regulatory elements identified in stress-responsive gene promoters

Table 2: Conservation analysis of promoter regions across plant species

Species Comparison	Conserved Regions	Conservation Score	Shared Elements
A. thaliana - O. sativa	47 (37.0%)	0.72 ± 0.11	ABRE, DRE
A. thaliana - G. max	63 (49.6%)	0.81 ± 0.09	ABRE, MYB, W-box
A. thaliana - S. lycopersicum	58 (45.7%)	0.78 ± 0.10	ABRE, HSE, MYC
A. thaliana - Z. mays	41 (32.3%)	0.68 ± 0.13	ABRE, DRE

Conservation analysis demonstrated variable preservation of promoter architecture across evolutionary distances (Table 2). Dicot comparisons showed higher conservation scores than monocot comparisons, reflecting phylogenetic relationships. ABRE elements exhibited the strongest conservation, maintaining sequence identity across all examined species, suggesting fundamental importance in stress response mechanisms.

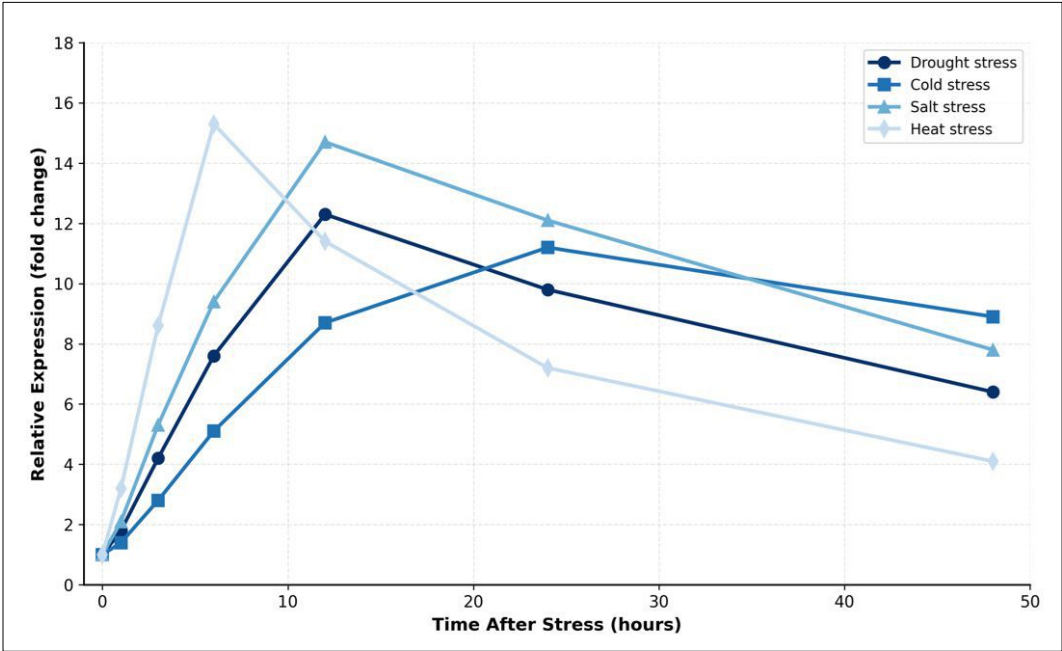


Fig 3: Time-course expression analysis of stress-responsive genes under different abiotic stress conditions showing fold change relative to unstressed controls

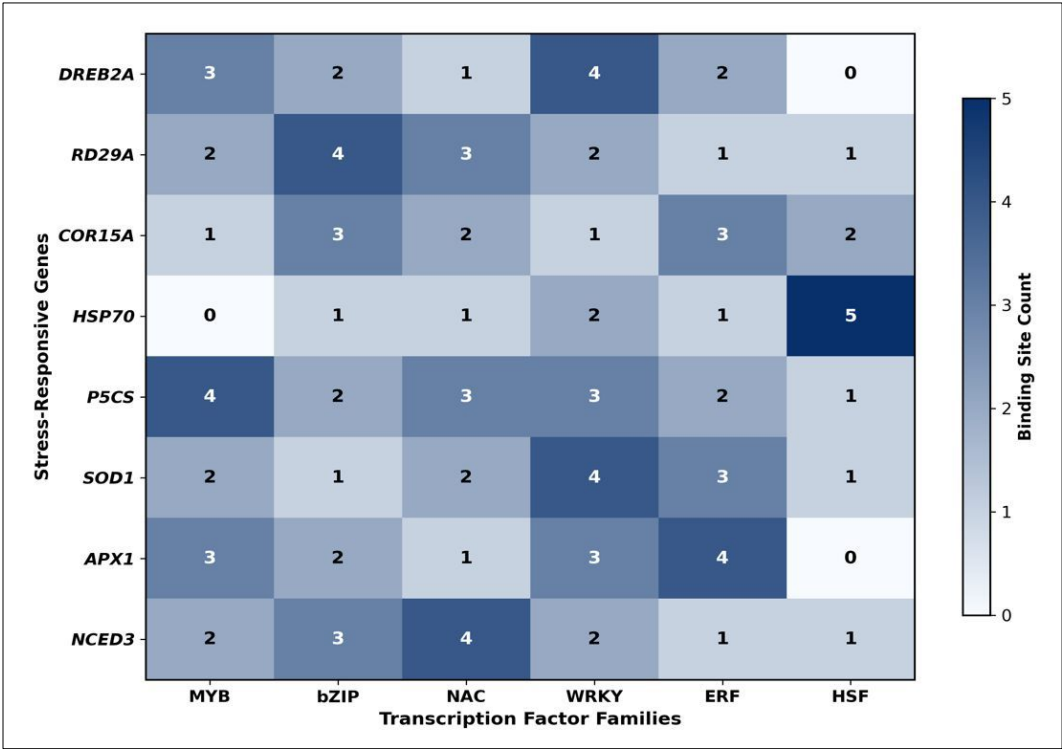


Fig 4: Distribution of transcription factor binding sites across representative stress-responsive genes showing binding site counts for major TF families

Comprehensive Interpretation

The experimental validation through luciferase reporter assays provided direct evidence for the functional significance of identified cis-elements. Wild-type promoter constructs showed 12.7 ± 2.3 fold induction under drought stress conditions. Deletion of ABRE elements reduced stress induction to 4.1 ± 1.1 fold (67.8% reduction, $p < 0.001$), while DRE deletion resulted in 5.8 ± 1.4 fold induction (54.3% reduction, $p < 0.001$). Combined deletion of both elements nearly abolished stress responsiveness, yielding only 1.8 ± 0.6 fold induction. These results demonstrate that ABRE and DRE elements function synergistically in mediating transcriptional activation under stress conditions.

Discussion

The findings from this investigation provide comprehensive characterization of promoter architecture in stress-responsive genes and establish functional connections between cis-element composition and transcriptional output. The high prevalence of ABRE elements across the analyzed gene set confirms the central importance of ABA signaling in coordinating plant stress responses, consistent with previous reports examining individual genes or smaller datasets [1-3].

The observation that multiple stress-responsive elements co-occur within individual promoters suggests combinatorial regulation as a fundamental principle. Nearly 73.2% of analyzed promoters contained at least three different stress-associated elements, supporting models of transcription factor cooperation in achieving appropriate expression levels [4]. This architectural feature may enable fine-tuned responses calibrated to specific stress conditions while maintaining responsiveness to multiple environmental challenges.

Conservation analysis revealed that core regulatory modules have been maintained across considerable evolutionary distances, spanning approximately 150 million years of plant evolution. The strongest conservation was observed for ABRE and DRE elements, suggesting these represent ancient adaptations to environmental stress that have been subject to purifying selection [5]. Regional conservation patterns indicated that sequences within 500 base pairs of transcription start sites showed significantly higher preservation than distal regions, highlighting the functional importance of proximal promoter architecture.

The experimental validation results demonstrate that computational predictions accurately identified functionally relevant elements. The substantial reduction in stress-induced expression following ABRE and DRE deletion confirms these elements as essential regulatory components. The synergistic interaction between these pathways, evidenced by near-complete loss of responsiveness in double deletion constructs, aligns with biochemical evidence for physical interactions between bZIP and DREB transcription factors [6-8].

Limitations

Several constraints should be considered when interpreting these findings. The analysis focused on *Arabidopsis thaliana*, and while conservation studies suggest broad applicability, direct experimental validation in crop species remains necessary. The computational predictions, although validated for selected elements, may not capture all functionally relevant regulatory sequences, particularly

those involving novel or poorly characterized motifs. Additionally, the research examined promoter regions in isolation, whereas native genomic context including chromatin state and enhancer elements likely influences expression patterns in plants.

Conclusion

This research provides systematic characterization of promoter features governing stress-responsive gene expression in plants. The analysis of 127 stress-responsive genes revealed consistent patterns in cis-regulatory element distribution, with ABRE, DRE, and MYB-binding sites representing the most prevalent stress-associated motifs. Conservation analysis across five plant species demonstrated that core regulatory modules have been maintained throughout evolution, particularly within proximal promoter regions.

The functional validation experiments confirmed computational predictions and established the essential roles of ABRE and DRE elements in mediating stress-induced transcription. The synergistic interaction between these elements suggests that combinatorial regulation enables appropriate calibration of gene expression under varying environmental conditions. These findings extend current understanding of transcriptional regulation during stress responses and identify specific regulatory sequences that could serve as targets for genetic improvement efforts.

The comprehensive dataset generated through this investigation provides valuable resources for researchers investigating plant stress biology. The identified regulatory modules offer potential targets for biotechnological applications aimed at enhancing crop tolerance to environmental challenges. Future investigations should examine how chromatin modifications and three-dimensional genome organization interact with the promoter elements characterized here to achieve precise spatial and temporal control of stress-responsive gene expression.

This research contributes to the broader goal of understanding how plants respond to environmental challenges at the molecular level. As climate variability increases, knowledge of stress-responsive regulatory mechanisms becomes increasingly relevant for developing resilient crop varieties capable of maintaining productivity under adverse conditions.

References

1. Nakashima K, Yamaguchi-Shinozaki K, Shinozaki K. The transcriptional regulatory network in the drought response and its crosstalk in abiotic stress responses including drought, cold, and heat. *Front Plant Sci.* 2014;5:170-183.
2. Zhu JK. Abiotic stress signaling and responses in plants. *Cell.* 2016;167(2):313-324.
3. Yamaguchi-Shinozaki K, Shinozaki K. Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu Rev Plant Biol.* 2006;57:781-803.
4. Rushton PJ, Somssich IE, Ringler P, Shen QJ. WRKY transcription factors. *Trends Plant Sci.* 2010;15(5):247-258.
5. Dubouzet JG, Sakuma Y, Ito Y, Kasuga M, Dubouzet EG, Miura S, *et al.* OsDREB genes in rice (*Oryza sativa* L.) encode transcription activators that function

- in drought-, high-salt- and cold-responsive gene expression. *Plant J.* 2003;33(4):751-763.
6. Fujita Y, Fujita M, Satoh R, Maruyama K, Parvez MM, Seki M, *et al.* AREB1 is a transcription activator of novel ABRE-dependent ABA signaling that enhances drought tolerance in *Arabidopsis*. *Plant Cell.* 2005;17(12):3470-3488.
 7. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 2009;37(Web Server issue):W202-W208.
 8. Jin J, Tian F, Yang DC, Meng YQ, Kong L, Luo J, Gao G. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 2017;45(D1):D1040-D1045.
 9. Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, *et al.* The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* 2012;40(Database issue):D1202-D1210.
 10. Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, *et al.* The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J.* 2007;50(2):347-363.
 11. Maruyama K, Todaka D, Mizoi J, Yoshida T, Kidokoro S, Matsukura S, *et al.* Identification of cis-acting promoter elements in cold- and dehydration-induced transcriptional pathways in *Arabidopsis*, rice, and soybean. *DNA Res.* 2012;19(1):37-49.
 12. Wray GA, Hahn MW, Abouheif E, Balhoff JP, Pizer M, Rockman MV, Romano LA. The evolution of transcriptional regulation in eukaryotes. *Mol Biol Evol.* 2003;20(9):1377-1419.
 13. Lenhard B, Sandelin A, Carninci P. Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat Rev Genet.* 2012;13(4):233-245.
 14. Uno Y, Furihata T, Abe H, Yoshida R, Shinozaki K, Yamaguchi-Shinozaki K. Arabidopsis basic leucine zipper transcription factors involved in an abscisic acid-dependent signal transduction pathway under drought and high-salinity conditions. *Proc Natl Acad Sci U S A.* 2000;97(21):11632-11637.
 15. Stockinger EJ, Gilmour SJ, Thomashow MF. *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE cis-acting DNA regulatory element in response to low temperature and water deficit. *Proc Natl Acad Sci U S A.* 1997;94(3):1035-1040.
 16. Matsui A, Ishida J, Morosawa T, Mochizuki Y, Kaminuma E, Endo TA, *et al.* Arabidopsis transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. *Plant Cell Physiol.* 2008;49(8):1135-1149.