THAI NGUYEN UNIVERSITY
UNIVERSITY OF EDUCATION

NGUYEN THU GIANG

FUNCTIONAL CHARACTERIZATION AND REGULATORY MECHANISMS OF GmDREB5 AND GmDREB7 TRANSCRIPTION FACTORS IN SALT STRESS RESPONSES OF TRANSGENIC TOBACCO AND SOYBEAN

Speciality: Genetics
Code: 9420121

DISSERTATION SUMMARY

THAI NGUYEN - 2026

This dissertation was completed at the Department of Genetics and Biotechnology, Faculty of Biology, Thai Nguyen University of Education.

Supervisors: 1. Assoc. Prof. Dr. Nguyen Huu Quan
2. Prof. Dr. Chu Hoang Mau

Reviewer 1: ……..

Reviewer 2: ……..

Reviewer 3: ……..

The dissertation will be defended in the university committee:
UNIVERSITY OF EDUCATION - THAI NGUYEN UNIVERSITY
***At* ……2026.**

The dissertation may be consulted at:

1. **National Library of Vietnam**
2. **Digital Center - Thai Nguyen University**
3. **Library of Thai Nguyen University of Education**

PUBLICATIONS RELATED TO THE DISSERTATION

1. Nguyen, G. T., Nguyen, Y. T. H., Nguyen, H. D., Chu, M. H., & Nguyen, Q. H. (2025), "The DREB7 transcription factor enhances salt tolerance in soybean plants under salt stress", Open life sciences, 20(1), 20251153. https://doi.org/10.1515/biol-2025-1153 (SCI-E, Q2).
2. Nguyen, Y. T. H., Nguyen, G. T., Nguyen, H. D., Vu, T. T. T., Nguyen, L. T. N., Nguyen, Q. H., & Chu, M. H. (2025), "The *GmDREB5* from soybean enhanced salt tolerance by upregulating the transcription levels of *NtP5CS* and *NtSODFe* in transgenic tobacco", Molecular biology reports, 52(1), 656. https://doi.org/10.1007/s11033-025-10767-x (SCI-E, Q2).
3. Nguyễn Thu Giang, Nguyễn Hữu Quân, Nguyễn Thị Hải Yến, Chu Hoàng Mậu (2025), "Evolution of the AP2 gene family and the DREB gene subfamily in *Arabidopsis thaliana* and soybean plants", *TNU Journal of Science and Technology*, 230(09), tr.398-405. https://doi.org/10.34238/tnu-jst.12362 (ACI).
4. Nguyen Thu Giang, Vu Trong Luong, Pham Thi Thu Hien, Nguyen Thi Phuong Thao, Nguyen Duc Hung, Nguyen Huu Quan, Vi Thi Xuan Thuy, Nguyen Thi Hai Yen, Chu Hoang Mau (2025), "Design of *GmDREB5* gene expression construction related to stress element response in soybean [Glycine max (L.) Merr.]", Kỷ yếu hội nghị sinh lý thực vật toàn quốc lần thứ 3 2025. Nxb Khoa học tự nhiên và Công nghệ, pp. 135-143.
5. Nguyễn Thu Giang, Nguyễn Thị Hải Yến, Vũ Thị Thu Thủy, Nguyễn Hữu Quân, Chu Hoàng Mậu (2026), "Characteristics of t1 generation soybean lines overexpressing the *GmDREB7* gene", *TNU Journal of Science and Technology*, 231(01), tr.19-25. https://doi.org/10.34238/tnu-jst.13248 (ACI).

INTRODUCTION

1. Rationale

Soybean (*Glycine max* (L.) Merrill) is an important crop with high economic and nutritional value. It also plays a significant role in improving soil fertility through its biological nitrogen fixation capacity. However, soybean yield and quality are often severely affected by adverse environmental factors, particularly salt and drought stresses. Under the impacts of climate change and long-term cultivation practices, soil salinization and water scarcity have become increasingly widespread, posing major challenges to agricultural production in general and soybean cultivation in particular.

Drought and salt tolerance in plants are complex traits regulated by intricate networks of functional and regulatory genes. Functional genes directly participate in stress responses by regulating osmotic adjustment, protecting cellular structures, or scavenging reactive oxygen species. These include genes involved in proline biosynthesis (*P5CS*, *P5CR*), root system development (*Expansin*), and heat shock proteins (*HSPs*). However, most of these genes are associated with only specific stress-response mechanisms. In contrast, transcription factor genes can simultaneously regulate multiple downstream target genes, thereby controlling entire stress-response networks in plants. Consequently, genes encoding transcription factors have emerged as promising targets for improving abiotic stress tolerance in crops through biotechnology approaches.

Among the transcription factor families identified in plants, the AP2/ERF (APETALA2/Ethylene Responsive Factor) family, particularly the DREB (Dehydration Responsive Element Binding) subfamily, has been recognized as a key regulator of plant responses to drought, salinity, and low-temperature stresses. DREB proteins contain a conserved AP2 domain that specifically recognizes and binds to DRE/CRT (Dehydration Responsive Element/C-repeat) cis-elements located in the promoters of target genes, thereby activating the expression of numerous stress-responsive genes. This ability to regulate multiple genes simultaneously makes DREB transcription factors important targets for research aimed at enhancing crop tolerance to unfavorable environmental conditions.

In soybean, numerous genes belonging to the DREB subfamily have been identified through genome-wide analyses and are predicted to

be involved in responses to abiotic stresses. However, the biological functions of most of these genes have not yet been fully elucidated, particularly their regulatory mechanisms underlying salt and drought stress responses in transgenic plants. Among them, *GmDREB5* and *GmDREB7* have been reported as potential regulators of abiotic stress responses based on their sequence characteristics and expression patterns. Nevertheless, direct functional evidence for their roles remains limited.

Based on these scientific and practical considerations, investigating the functions of DREB transcription factor genes in soybean is of great importance for elucidating the molecular mechanisms underlying salt and drought tolerance. Such studies also provide a foundation for the application of genetic engineering technologies in developing crop varieties with enhanced adaptation to adverse environmental conditions. Therefore, this dissertation was conducted under the title:**"Functional characterization and regulatory mechanisms of GmDREB5 and GmDREB7 transcription factors in salt stress responses of transgenic tobacco and soybean"**.

2. Research objectives

To elucidate the regulatory roles of the genes encoding DREB5 and DREB7 proteins in controlling the expression of selected target genes under abiotic stress conditions.

To generate and evaluate transgenic soybean lines overexpressing GmDREB7 with enhanced salt stress tolerance compared with non-transgenic control plants.

3. Research Contents

3.1. Investigation of the molecular evolution of the AP2 gene family in *Arabidopsis thaliana* and soybean

i) Bioinformatics approaches were employed to identify and retrieve sequences of genes belonging to the AP2 family, particularly the DREB subfamily, from *A. thaliana* and soybean using the NCBI database.

ii) The number of *AP2* family genes and *DREB* subfamily genes, as well as the chromosomal locations and copy numbers of each *DREB* gene, were determined in the genomes of *A. thaliana* and soybean.

iii) Phylogenetic trees were constructed for the *AP2* gene family, the *DREB* subfamily, and the AP2 conserved domain in *A. thaliana* and soybean.

3.2. Functional analysis of the DREB5 transcription factor gene in soybean

i) Design of the *GmDREB5* (*Glycine max DREB5*) gene construct and expression vector.
ii) Genetic transformation of tobacco plants with the *GmDREB5* construct and generation of *GmDREB5* transgenic lines.
iii) Analysis of the relationship between *GmDREB5* expression and the transcriptional levels of selected target genes in transgenic tobacco plants.
iv) Analysis of the interactions between the DREB5 protein and the promoters of selected target genes.

3.3. Investigation of the role of the *GmDREB7* transcription factor gene in soybean

i) Genetic transformation of soybean with the *GmDREB7* (*Glycine max DREB7*) construct and generation of transgenic soybean lines.
ii) Molecular characterization of transgenic soybean plants using real-time RT-PCR analysis.
iii) *In silico* analysis of the molecular interaction between the DREB7 protein and the promoter of the *GmP5CS* gene.
iv) Evaluation of selected promising *GmDREB7* transgenic soybean lines for drought and salt tolerance.

4. Novel contributions of the dissertation

4.1. The dissertation identified the regulatory role of the *GmDREB5* gene in response to salt stress. *GmDREB5* is capable of activating the expression of the target genes *NtP5CS* and *NtSODFe* in tobacco plants, thereby enhancing cellular protection mechanisms through the accumulation of osmoprotectants and the increased activity of reactive oxygen species (ROS)-scavenging enzymes.
4.2. The regulatory function of *GmDREB7* in salt stress responses was clearly demonstrated. Overexpression of *GmDREB7* in transgenic soybean plants enhanced the transcription of *GmP5CS*, leading to increased proline accumulation and improved adaptation to salt stress. Four transgenic soybean lines (*TG1-2, TG1-5, TG1-7,* and *TG1-10*) exhibiting markedly enhanced tolerance to both salinity and drought stresses were successfully selected in comparison with wild-type (WT) plants.
4.3 The dissertation elucidated the molecular basis of the interaction between the AP2 domain of DREB proteins and the promoters of the target genes (*NtP5CS, NtSODFe,* and *GmP5CS*) through *in silico* molecular docking analysis. These findings provided further evidence for the specific binding of DREB transcription factors to stress-responsive *cis*-regulatory elements.

5. Scientific and practical significance of the dissertation

The results of this study provided new insights into the regulatory mechanisms governing salt and drought stress responses in soybean, highlighting the roles of transcription factors from the DREB subfamily. The identification of the regulatory functions of *GmDREB5* and *GmDREB7* provided experimental evidence linking the activity of these genes to osmolyte biosynthesis and reactive oxygen species (ROS) scavenging enzymes, thereby elucidating the molecular basis of salt tolerance in soybean. These results expand current knowledge of the gene regulatory networks underlying the physiological adaptation of plants to adverse environmental stresses.

The study provided a scientific foundation for the application of genetic engineering in the development of salt and drought tolerant soybean varieties, thereby opening a promising avenue for enhancing crop stress tolerance and promoting agricultural adaptation to climate change.

The findings published in national and international scientific journals serve as valuable resources for research and teaching in the fields of biology and plant biotechnology.

6. Structure of the dissertation

This dissertation consists of 162 pages (including appendices). The structure of the dissertation is as follows: Introduction (4 pages); Chapter 1: Literature Review (40 pages); Chapter 2: Materials and Methods (12 pages); Chapter 3: Results and Discussion (53 pages); Conclusions and Recommendations (2 pages); Publications Related to the Dissertation (1 page); References (15 pages); and Appendices (20 pages). This dissertation contains 6 tables, 23 figures, 13 appendices, and 134 references.

CHAPTER 1. LITERATURE REVIEW

The dissertation reviewed and synthesized 131 English-language references on three main topics: 1) Characteristics of the AP2 transcription factor family and the DREB subfamily; 2) Functional characterization of soybean genes using model plants; 3) Roles and applications of transcription factor-encoding genes in improving drought and salinity tolerance in soybean.

1.1. Characteristics of the AP2 transcription factor family and the DREB subfamily

Transcription factors (TFs) play pivotal roles in orchestrating and coordinating plant responses to abiotic stresses. Through their ability to specifically recognize and bind cis-regulatory elements located in the promoter regions of target genes, TFs regulate gene transcription, thereby contributing to the maintenance of osmotic homeostasis, preservation of cellular structural integrity, and mitigation of damage caused by the accumulation of reactive oxygen species. Among the transcription factor families identified in plants, the AP2 family is distinguished by its large size and extensive functional diversity. Members of this family are characterized by the highly conserved AP2 DNA-binding domain, together with specific functional regions that enable flexible interactions with various cis-elements and participation in the regulation of a wide range of physiological processes.

Within the AP2 family, the DREB subfamily is regarded as one of the most important groups directly involved in plant responses to drought and salinity stresses, primarily through abscisic acid (ABA) independent signaling pathways. Phylogenetic and functional studies across diverse plant species have demonstrated that *DREB* genes, particularly those belonging to the A-2 and A-5 groups, can simultaneously activate multiple stress-responsive genes associated with osmotic adjustment, redox homeostasis, and protein stabilization under adverse environmental conditions.

In soybean, the large number and remarkable diversity of AP2/ERF family genes, including members of the DREB subfamily, reflect the consequences of whole-genome duplication events during evolution as well as the species' high ecological adaptability. Several *DREB* genes, such as *GmDREB1*, *GmDREB2*, *GmDREB3* and *GmDREB6*, have been reported to play important roles in drought tolerance, salinity tolerance, and responses to high temperature stress. However, the biological functions of many other DREB genes, including *GmDREB5* and *GmDREB7*, remain insufficiently characterized. Therefore, investigating these two genes is of considerable significance not only for elucidating the regulatory mechanisms underlying abiotic stress responses in soybean but also for providing a scientific basis for the development of drought and salinity tolerant soybean cultivars, thereby supporting sustainable agricultural production under changing climatic conditions.

1.2. Functional characterization of soybean genes using model plants

Recent studies have demonstrated that model plant systems play a crucial role in the functional characterization of plant genes. Among these systems, tobacco (*Nicotiana tabacum*) has been widely recognized as an efficient heterologous platform for investigating soybean gene function due to its high transformation efficiency, strong regenerative capacity, and substantial biomass production. Using this model system, numerous soybean genes have been functionally analyzed and shown to participate in plant responses to various abiotic stresses. In particular, genes encoding transcription factors of the AP2/ERF family, including members of the DREB subfamily, have been identified as central regulatory components in stress-response networks through their control of the expression of multiple downstream target genes.

Studies on transgenic tobacco plants expressing soybean DREB genes, such as *GmDREB2* and *GmDREB6*, have shown that the expression of these genes enhances drought and salinity tolerance by modulating proline accumulation and activating the expression of stress responsive genes, including *NtP5CS* and *NtCLC*. In addition, other members of the DREB family have been reported to regulate stress responses through distinct mechanisms. For example, *GmDREB7* has been shown to function as a negative regulator of certain salt stress responsive genes in transgenic tobacco plants, highlighting the functional diversity and regulatory complexity of this gene family.

Despite these advances, current knowledge regarding the functions of soybean DREB genes remains limited. Many members of this gene family have not yet been comprehensively characterized with respect to their biological roles and regulatory mechanisms within plant stress response networks. In particular, *GmDREB5*, a member of the soybean DREB subfamily, has not been systematically investigated regarding its biological function or its potential role in regulating stress-responsive genes in plants. Therefore, further functional characterization of *GmDREB5* using model plant systems, especially transgenic tobacco, is necessary to elucidate its role in plant abiotic stress responses.

1.3. Roles and applications of transcription factor encoding genes in improving drought and salinity tolerance in soybean

Abiotic stresses, particularly drought and salinity, are among the most significant environmental factors limiting plant growth and crop productivity. These stresses reduce cellular water potential, induce stomatal closure, inhibit photosynthesis, disrupt ion homeostasis, and

promote the accumulation of reactive oxygen species, including free radicals such as superoxide (O_2^-) and hydroxyl radicals (•OH), as well as non-radical reactive molecules such as hydrogen peroxide (H_2O_2).

Numerous studies have demonstrated that soybean can adapt to adverse environmental conditions through a variety of physiological and biochemical mechanisms. These include osmotic adjustment via the accumulation of compatible solutes such as proline, glycine betaine, and soluble sugars; enhancement of antioxidant enzyme activities; and modulation of hormonal signaling pathways and stress responsive gene expression. Within this regulatory network, genes encoding transcription factors, particularly members of the DREB subfamily within the AP2/ERF family, are considered key regulators due to their ability to coordinately control the expression of multiple stress related target genes. Transgenic studies have shown that the overexpression of certain DREB genes can enhance drought and salinity tolerance by regulating a range of physiological and biochemical processes associated with stress adaptation.

Despite these advances, the biological functions of many soybean DREB family members remain insufficiently characterized. In addition to experimental approaches, bioinformatics-based methods, particularly molecular interaction analyses between the AP2 DNA-binding domain of DREB proteins and the promoter regions of target genes, are increasingly being employed to predict regulatory mechanisms and identify genes with potential functional significance. Therefore, further investigation of DREB genes in soybean is necessary to elucidate the molecular mechanisms underlying abiotic stress responses and to provide a scientific foundation for the development of soybean cultivars with enhanced drought and salinity tolerance.

CHAPTER 2. MATERIALS AND METHODS

2.1. Research materials

Plant materials: The tobacco cultivar K326 (*Nicotiana tabacum* L.) grown under in vitro conditions was obtained from the Plant Cell Biotechnology Laboratory, Faculty of Biology, Thai Nguyen University of Education. The soybean cultivar DT26, which is widely cultivated in Vietnam, was obtained from the Legume Research and Development Center, Vietnam Academy of Agricultural Sciences.

Bacterial strains and vectors: *Escherichia coli* strains used for plasmid propagation and the Agrobacterium tumefaciens strain AGL1 carrying the *pBI121_GmDREB7* expression construct were obtained from a project funded by the Ministry of Education and Training of Vietnam (Grant No. B2023-TNA-26) and were maintained at the Plant Cell Biotechnology Laboratory, Faculty of Biology, Thai Nguyen University of Education.

The vectors employed in this study included the cloning vector pTWIST (Twist Bioscience, USA), the plant expression vector pFGC5941 (NovoPro Bioscience, USA) and the plant expression vector pBI121 (Bioactiva Diagnostica, Germany), all of which contain the CaMV35S promoter.

2.2. Chemicals and equipment

Chemicals: Molecular biology kits, restriction endonucleases and DNA ligases used in this study were obtained from reputable commercial suppliers. Chemicals including peptone, yeast extract, agarose, sucrose, glucose, tryptone, X-gal, KCl, Tris–HCl, EDTA, NaOH, $MgSO_4$, $MgCl_2$, glycerol, and $CaCl_2$, as well as selective antibiotics such as kanamycin, rifamycin, cefotaxime, and carbenicillin, were purchased from established manufacturers.

Equipment: The equipment used in this study included a PCR System 9700 thermal cycler, electrophoresis apparatus, gel documentation system, vortex mixer, centrifuge, electroporator, nucleic acid quantification instrument, and other modern laboratory equipment supporting molecular biology and plant tissue culture research.

2.3. Research methods

The study was conducted using four major methodological approaches: (1) Bioinformatics analyses; (2) Gene transfer vector construction and heterologous expression in tobacco and soybean plants; (3) Gene expression analysis and molecular docking studies and (4) Data analysis and statistical evaluation.

2.3.1. Bioinformatics approaches:

The bioinformatics analyses included:

(i) The collection and curation of sequence data for members of the *AP2* gene family and the *DREB* subfamily; (ii) Characterization and molecular evolutionary analyses of *AP2* family genes and *DREB*

subfamily genes and (iii) Evolutionary analysis of the AP2 domain among DREB proteins.

2.3.2. Gene transfer vector construction and expression in tobacco and soybean plants

The methods in this group included gene and vector design, generation of recombinant *A. tumefaciens* strains and agrobacterium mediated transformation of tobacco and soybean plants.

2.3.3. Gene expression and functional analyses

This group included salt stress treatment of transgenic lines, transgene detection, transgene expression analysis, proline quantification, functional characterization of the transgene and molecular docking studies.

2.3.4. Data analysis and statistical methods

Bioinformatics analyses were conducted using genome analysis software, including MEGA11, BioEdit and BLAST.

Statistical analyses were performed using graphPad prism (GraphPad Software, San Diego, CA, USA). Analysis of variance (ANOVA) followed by tukey's multiple comparison test was used to evaluate differences among groups. Statistical significance was determined at p-values of < 0.05, < 0.01, and < 0.001, depending on the parameter analyzed.

2.4. Research site and completion of the dissertation

The dissertation was completed at the Faculty of Biology, University of Education, Thai Nguyen University.

CHAPTER 3. RESULTS AND DISCUSSION

3.1. MOLECULAR EVOLUTION OF THE AP2 GENE FAMILY IN *A.THALIANA* AND SOYBEAN

3.1.1. Evolution of the AP2 gene family in *A. thaliana* and soybean

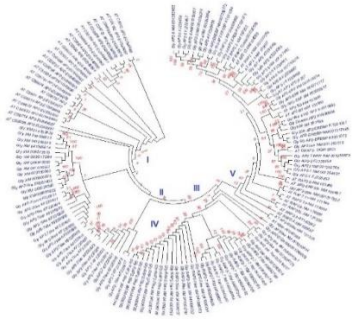


Figure 3.1 Phylogenetic tree and molecular evolutionary relationships of AP2 gene family members in *A. thaliana* and soybean, constructed using the Maximum Likelihood (ML) method with the JTT matrix-based model in MEGA.

Based on BLAST analysis against the NCBI database, a total of 143 AP2 family gene sequences were identified and collected. A phylogenetic tree was constructed using the Maximum Likelihood (ML) method with the JTT amino acid substitution model implemented in MEGA software. The phylogenetic analysis revealed that the AP2 sequences from *A. thaliana* and soybean were classified into five major groups, designated as groups I-V (Hình 3.1).

The phylogenetic analysis showed that group I comprised 24 AP2 sequences specific to *A. thaliana*, representing a distinct evolutionary lineage associated with the evolutionary history of this species. In contrast, groups II, III, and IV contained 34, 11 and 31 soybean AP2 sequences, respectively, forming monophyletic clades characteristic of soybean. Notably, group V was identified as a paraphyletic group consisting of a total of 53 AP2 sequences, including 18 sequences from *A. thaliana* and 35 sequences from soybean. The coexistence of sequences from both species within the same phylogenetic group suggests a high degree of structural and functional conservation of certain AP2 genes throughout the evolutionary divergence of these two species.

3.1.2. Evolution of the *DREB* gene subfamily in *A. thaliana* and soybean

In this study, 50 DREB protein sequences from *A. thaliana* were used as query sequences to identify homologous genes in the soybean

genome. The analysis identified 78 protein sequences corresponding to 18 distinct *DREB* genes in soybean. In *A. thaliana*, *DREB* genes were distributed across all five chromosomes, with a total of 12 *DREB* genes identified. Among them, chromosome 1 contained the highest number of *DREB* sequences (15 sequences), whereas chromosome 3 harbored only five sequences, representing the lowest number among the five chromosomes.

3.1.3. Characteristics of the AP2 domain in DREB proteins of *A. thaliana* and soybean

The resulting phylogenetic tree, constructed from 128 AP2 domain amino acid sequences, was divided into 16 distinct groups (Figure 3.3). Among these, nine groups were identified as monophyletic clades (groups 1, 2, 3, 4, 5, 7, 12, 13, and 16), whereas seven groups were classified as paraphyletic clades (groups 6, 8, 9, 10, 11, 14, and 15). The monophyletic groups generally contained a relatively small number of sequences, ranging from 2 to 8, while the paraphyletic groups comprised larger numbers of sequences, ranging from 4 to 28. This distribution pattern reflects both the conservation and diversification of the AP2 domain within the DREB subfamily and suggests the involvement of gene duplication and functional divergence events in the evolutionary history of this gene family.

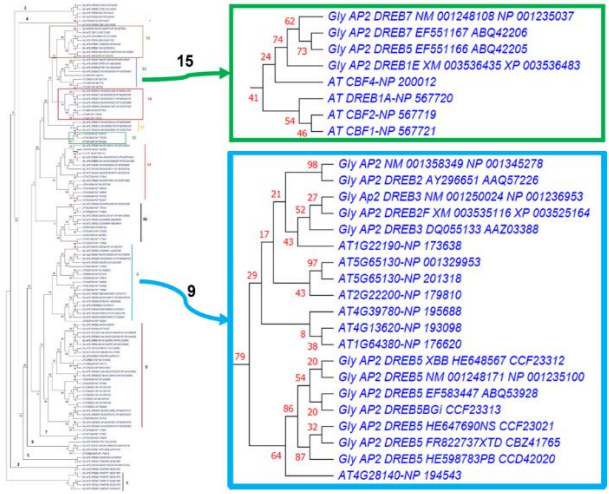


Figure 3.3. Phylogenetic tree of the AP2 domains from the DREB5 and DREB7 protein subfamilies in *A. thaliana* and soybean, constructed using the Maximum Likelihood method. (Gly: Glycine max = Gm; AT: *A. thaliana)*

Thus, compared with previous studies, the present study not only reinforces the central role of whole genome duplication (WGD) events in the expansion of the AP2 gene family in soybean but also provides further insights into the patterns of evolutionary conservation and divergence of this gene family at multiple levels, including the AP2 family, the DREB subfamily and the AP2 functional domain. These findings contribute to a deeper understanding of the molecular evolution of AP2 transcription factors and provide an important scientific basis for the identification and utilization of candidate genes for functional characterization, as well as for developing strategies to improve drought and salinity tolerance in crop plants.

3.2. FUNCTIONAL CHARACTERIZATION OF THE GENE ENCODING THE DREB5 TRANSCRIPTION FACTOR IN SOYBEAN

3.2.1. Characteristics of *GmDREB5* gene and expression vector carrying *GmDREB5* gene

After completion of the design process, the artificially synthesized *GmDREB5* gene, consisting of 1,006 nucleotides (Figure 3.4), was cloned into the pTwist cloning vector to generate the recombinant vector *pTWIST_GmDREB5-syn*. This recombinant construct was subsequently transformed into E. coli for propagation and maintenance of the target gene. Positive bacterial colonies harboring plasmids containing *GmDREB5-syn* were screened and selected according to appropriate criteria. Subsequently, the *pTWIST_GmDREB5-syn* plasmids from the selected bacterial clones were amplified, isolated, and purified for further analyses.

Figure 3.4. Nucleotide sequence of *GmDREB5* containing *Asc*I and *BamH*I restriction sites and the c-Myc and KDEL coding sequences.

The purified *GmDREB5* gene fragment was ligated into the pFGC5941 expression vector through a ligation reaction catalyzed by T4 DNA ligase, generating the recombinant vector *pFGC_GmDREB5* (Figure 3.6). To identify bacterial clones harboring the recombinant *pFGC_GmDREB5* vector, colony PCR was performed using two primer pairs specific for the *GmDREB5* and Bar genes. As shown in Figure 3.7A, lane 1 displayed a DNA band of approximately 856 bp, consistent with the expected size of the *GmDREB5* gene, whereas no amplification product was detected in lane 2, indicating the absence of the target gene in this clone.

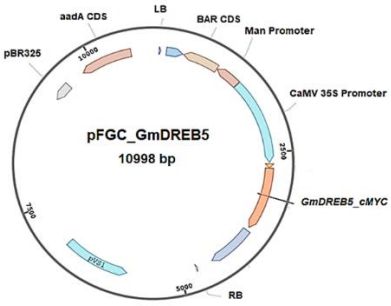


Figure 3.6. Structure of the pFGC_GmDREB5 expression vector. aadA CDS: coding sequence of aminoglycoside phosphotransferase; LB: left border

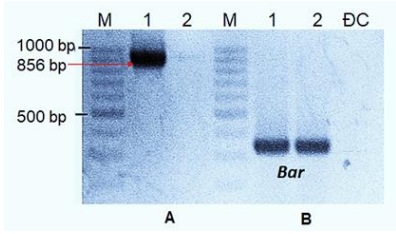


Figure 3.7. Agarose gel electrophoresis of colony-PCR products used for colony screening.
A: PCR products amplified using

repeat derived from C58 T-DNA; BAR CDS: coding sequence of phosphinothricin acetyltransferase; Man promoter: mannopine synthase promoter; CaMV 35S promoter: cauliflower mosaic virus 35S promoter; GmDREB5-cMyc: *GmDREB5* gene region fused with the c-Myc coding sequence; RB: right border repeat derived from C58 T-DNA.

the primer pair *GmDREB5-F/MYC-KDEL-R*; B: PCR products amplified using the primer pair *BarR3-F/BarR1-R*. M: DNA marker; Lane 1: selected recombinant *A. tumefaciens* colony; Lane 2: pFGC5941 plasmid; ĐC: negative control.

3.2.2. Transformation of pFGC_GmDREB5 construct and selection of tobacco lines carrying *GmDREB5*

The detailed results of each transformation experiment are presented in Table 3.2. Across three independent experiments, a total of 90 tobacco leaf explants were used for transformation with the *GmDREB5* gene construct. Following the co-cultivation and shoot regeneration stages, 70 explants remained viable, producing 162 shoot clusters. Among these, 35 shoots successfully developed roots on the selective medium, and a total of 23 complete plantlets were obtained and subsequently transferred to substrate in the greenhouse for growth monitoring and further molecular analyses.

Table 3.2. Results of transformation of pFGC_GmDREB5 construct into tobacco

Experimental and control groups		**Count of transformed explants**	**Count of viable explants**	**Count of shoot clusters formed**	**Count of selected surviving shoots**	**Count of selected rooted shoots**	**Count of *in vitro* plants generated**
Experiment	1	30	25	62	26	12	7
	2	30	22	42	15	8	6
	3	30	23	58	11	15	10
Total		***90***	***70***	***162***	***52***	***35***	***23***
Control 1		30	0	0	0	0	0
Control 0		30	22	45	88	35	12

Among the transformed tobacco plants, 18 TG0 transgenic plants maintained in vitro were selected for the establishment of transgenic lines. Three transgenic lines (TG02, TG08 and TG15),

together with the wild-type (WT) plants exhibiting normal growth, were selected for subsequent studies.

The results of RT-PCR and Western blot analyses demonstrated that the *GmDREB5* gene isolated from soybean was not only stably integrated into the genome of tobacco cultivar K326 but was also efficiently transcribed and translated into the corresponding recombinant protein in the TG0 generation. These findings provide reliable evidence for the functional expression of *GmDREB5* in a model plant system and establish a solid foundation for subsequent studies aimed at elucidating the biological role of this gene in enhancing drought and salt tolerance in transgenic plants.

3.2.3. Analysis of the relationship between *GmDREB5* and the transcript levels of selected target genes in tobacco plants

3.2.3.1. Effect of GmDREB5 expression on the transcript levels of NtP5CS and NtSODFe in transgenic tobacco plants

The results of comparing the transcription levels of the *GmDREB5* transgene among transformed tobacco lines and between transformed lines and WT plants under no-stress and salt stress conditions are presented in Figure 3.10. The graph in Figure 3.10 shows that under both regular watering and salt stress conditions, the transcription levels in the TG08 and TG015 lines were higher than those in the TG02 line and WT plants. Under salt stress conditions, the transcription levels in the TG08 and TG015 lines increased significantly compared to wild-type (WT) plants, with 87.31-fold (TG08) and 39.07-fold (TG015) increases.

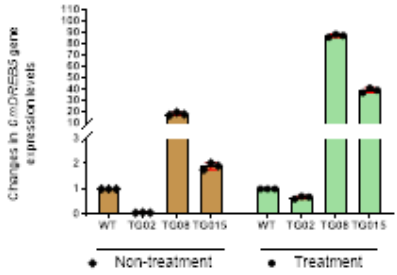
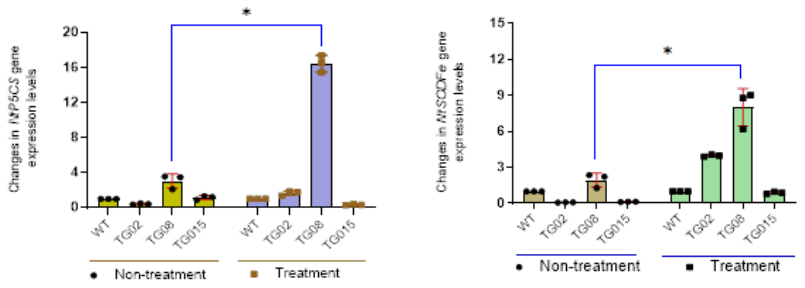


Figure 3.10. Changes in the transcription levels of the *GmDREB5* gene among transgenic tobacco lines and WT plants under no-stress and salt stress conditions.



A **B**

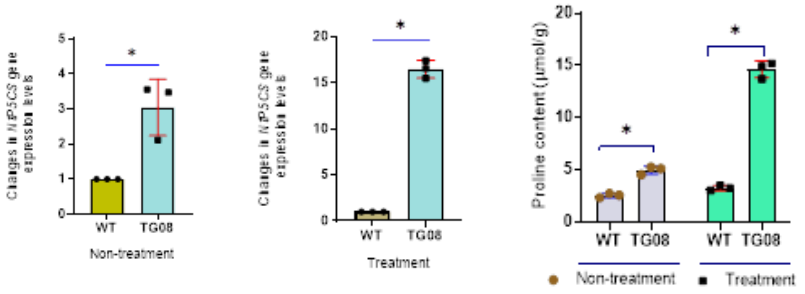
Figure 3.11. Changes in the expression levels of *NtP5CS* (A) and *NtSODFe* (B) genes in *GmDREB5* transgenic tobacco lines under no-stress and salt stress conditions were assessed using qRT-PCR.

The relationship between the *GmDREB5* transgene and the expression levels of the *NtP5CS* and *NtSODFe* genes in transgenic tobacco lines and WT plants under two conditions, regular watering and salt stress, is presented in Figure 3.11. In Figure 3.11A, the *NtP5CS* gene in the TG08 transgenic line showed significantly higher transcription levels than in WT plants, with 3.05-fold (non stress) and 16.47fold (salt stress) increases. Under salt stress, the *NtP5CS* expression in TG08 was 5.41 times higher than under non stress conditions.

In Figure 3.11B, the *NtSODFe* gene in both TG02 and TG08 lines exhibited significantly higher transcription levels under salt stress than under non-stress conditions, with increases of 4.12 fold (TG08) and 51.60 fold (TG02) ($P < 0.05$). Furthermore, the transcription levels of *NtSODFe* in TG02 and TG08 were also significantly higher than in WT plants, with increases of 1.95 fold (non-stress) and 8.01 fold (salt stress) for TG08 and 0.08 fold (non-stress) and 3.97 fold (salt stress) for TG02 ($P < 0.05$).

Thus, the expression levels of the *NtP5CS* and *NtSODFe* genes in the *GmDREB5* transgenic tobacco lines TG02 and TG08 were higher than those in WT plants under salt stress conditions. This result indicates that the soybean *GmDREB5* gene could enhance the expression levels of the *NtP5CS* and *NtSODFe* genes in transgenic tobacco plants

3.2.3.2. Relationship between NtP5CS transcript levels and proline content in transgenic tobacco lines



A **B** **D**

Figure 3.12. Changes in the transcription levels of the *NtP5CS* gene (A; B) and proline content (D) among TG08 transgenic tobacco line and WT plants under no-stress and salt stress conditions ($P<0.05$).

Figure 3.12 illustrates the relationship between the transcription level of the *NtP5CS* gene and proline accumulation in the transgenic tobacco line TG08 compared with WT plants under normal growth conditions and salt stress. Upon salt treatment, this difference became markedly more pronounced, with *NtP5CS* expression in TG08 being 16.47 fold higher than that in WT plants (Figure 3.12B). Consistent with the increased transcription of *NtP5CS*, the proline content in TG08 also showed a corresponding increase. Specifically, under non-stress conditions, the proline content in TG08 was 1.96 fold higher than that in WT plants, whereas under salt stress, this difference increased to 4.50 fold (Figure 3.12D)

3.2.4. Molecular docking between the AP2 domain of DREB5 and the promoter regions of selected target genes

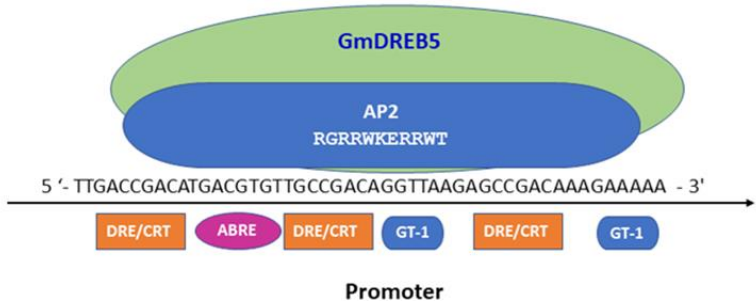


Figure 3.13. Molecular docking analysis of the AP2 domain of GmDREB5 with *cis*-regulatory motifs in the promoter regions of the *NtP5CS* and *NtSODFe* genes.

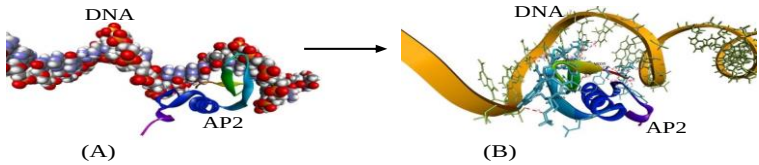


Figure 3.14. AP2-promoter docking models and the hydrogen-bond (H-bond) network between AP2 and target DNA.

Molecular docking analyses provided additional evidence supporting the proposed regulatory mechanism at the molecular level. The AP2 domain of GmDREB5 was predicted to interact directly with key *cis*-regulatory elements within the promoter regions of *NtP5CS* and *NtSODFe*, including the GT-1 and DRE/CRT motifs. Conserved amino acid residues within the AP2 domain were involved in the formation of stable hydrogen bonds with DNA, suggesting a specific and biologically relevant DNA-binding capability.

Docking property analysis showed that the AP2-P5CS promoter complex displayed more favorable characteristics than the AP2-SODFe promoter complex, including better constraint satisfaction, reduced solvent influence, more stable electrostatic interactions, and stronger overall binding affinity. These *in silico* findings support experimental results in which *GmDREB5* overexpression enhanced *NtP5CS* and *NtSODFe* expression in transgenic tobacco plants.

3.3. FUNCTIONAL ROLE OF THE DREB7 TRANSCRIPTION FACTOR GENE IN SOYBEAN

3.3.1. Genetic transformation of the pBI121_GmDREB7 construct into soybean and creation of transgenic plants

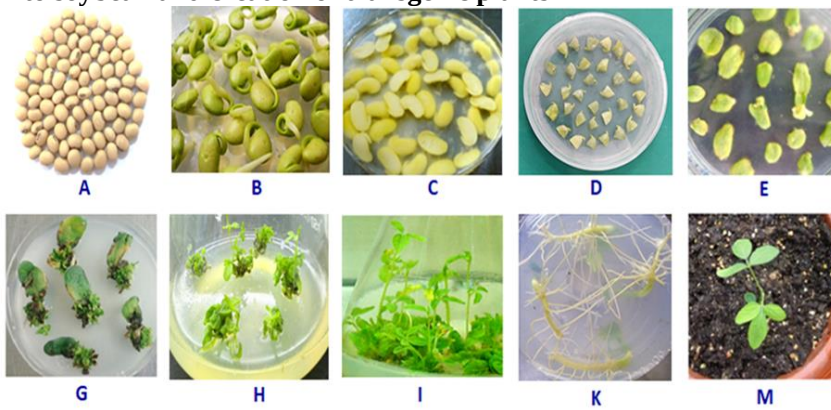


Figure 3.16. Transformation of pBI121_DREB7 construct and regeneration of transgenic soybeans.

The designed gene construct consisted of a 624bp coding sequence, a c-MYC epitope tag coding sequence (33 bp), a KDEL signal peptide sequence (12 bp), and recognition sites for the restriction enzymes *Xba*I, *BamH*I, *Xma*I and *Sac*I. Upon completion, the *GmDREB7* gene construct was introduced into soybean via *A. tumefaciens* mediated transformation using recombinant *A. tumefaciens* strains and cotyledonary node explants as the transformation material. This approach was employed to generate transgenic soybean lines for subsequent functional analyses.

Monitoring the growth and development of nine transgenic plants in the TG0 generation with positive RT-PCR results showed that seven plants flowered and produced fruit (TG0-2, TG0-3, TG0-4, TG0-5, TG0-6, TG0-7, TG0-10). However, only the fruits of four TG0 plants, TG0-2, TG0-5, TG0-7, and TG0-10, produced seeds; the remaining plants, TG0-3, TG0-4 and TG0-6, produced empty fruits and did not produce seeds. The seeds of four TG0 plants (TG0-2, TG0-5, TG0-7, and TG0-10) were germinated and grown in a greenhouse to produce four TG1 transgenic lines, namely TG1-2, TG1-5, TG1-7, and TG1-10.

3.3.2. Results of the analysis of transgenic soybean plants

3.3.2.1. transgenic soybean plants in TG1 generation and WT plants under salt stress

The results of observing the morphology of the plants after two days of each salinity treatment showed that, in the first treatment, the morphology of WT plants and transgenic lines did not show any morphological expression due to the impact of 150 mM NaCl. However, the second treatment with 250 mM NaCl affected the morphology of transgenic soybean plants and WT plants and the most substantial effect was after the third treatment. WT plants and TG1-7 soybean lines showed the most evident signs of water loss due to salinity stress.

3.3.2.2. Analysis of GmDREB7 transcript levels in transgenic soybean lines and wild-type plants after 7 days of NaCl stress

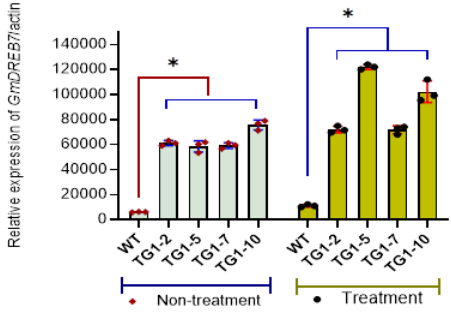


Figure 3.19. The diagram compares the transcription levels of the *GmDREB7* gene by Real-time RT-PCR in WT and TG1 generation transgenic soybean lines under untreated and salt-treated conditions.

WT plants and TG1 transgenic soybean lines after two days of the second treatment with 50 mL NaCl 250 mM were selected to analyze the *GmDREB7* gene expression level of the transgenic lines and WT by RT-PCR. Leaves of WT and transgenic soybeans in the TG1 generation were used to extract total RNA, generate cDNA, and analyze the transcription level of the *GmDREB7* of WT plants and four transgenic lines by Real-time RT-PCR; the results are shown in the graph in Figure 3.19. In Figure 3.19, all transgenic soybean lines exhibited higher transcription levels of the *GmDREB7* gene compared to WT plants. Under salt stress conditions, after salinity treatment with 250 mM NaCl, the transgenic soybean lines TG1-5 and TG1-10 had the strongest and highest transcription levels compared to those under non-salinity treatment conditions ($P<0.05$).

3.3.2.3. Analysis of changes in proline content in transgenic soybean lines and wild-type plants under salt stress conditions

The amino acid proline is an osmolyte; increasing the proline content will increase the water retention capacity of the cell, and the

plant will be resistant to salt stress. Therefore, this study continues to investigate the relationship between the overexpression of the *GmDREB7* gene and the transcription of the gene encoding the key enzyme P5CS, an enzyme involved in proline synthesis, in transgenic lines and WT plants under salt stress and normal conditions. The results of analyzing the transcription level of the *GmP5CS* gene of the *GmDREB7* transgenic lines and WT plants showed that the two transgenic lines, TG1-5 and TG1-10, had higher expression levels than WT plants and higher than in non-salt treated conditions with P<0.05 (Figure 3.20).

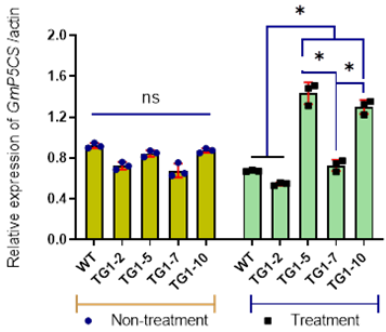


Figure 3.20. The diagram compares *GmP5CS* gene transcription levels in four TG1 generation transgenic soybean lines and wild-type plants using RT-PCR.

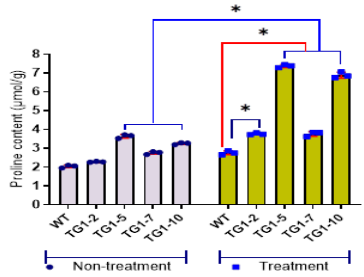
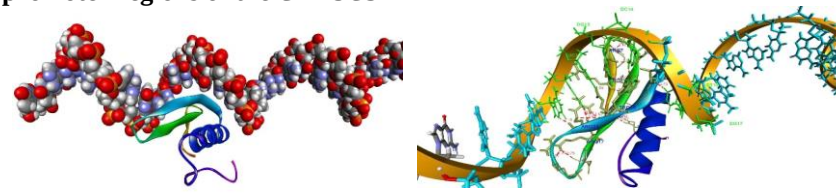


Figure 3.21. The diagram compares the proline content of the TG1 generation transgenic soybean lines with that of the WT plants.

The results in Figure 3.21 showed that proline content in transgenic lines was higher than in WT plants and increased from 136.59% to 267.03%. Among the transgenic soybean lines, lines TG1-5 and TG1-10 had the highest proline contents, 7.37 ± 0.05 (µmol/g) and 6.89 ± 0.09 (µmol/g), respectively (P<0.05); and the proline content of the two this transgenic lines under salt stress conditions was both higher than under untreated conditions (Figure 5B, C). These results suggested that the *GmDREB7* transgenic soybean lines may have had higher salt tolerance than WT plants, and lines TG1-5 and TG1-10 had the highest salt tolerance among the transgenic soybean lines in this study.

3.3.3. Docking between the AP2 region of the DREB7 protein and the promoter regions of the *GmP5CS*



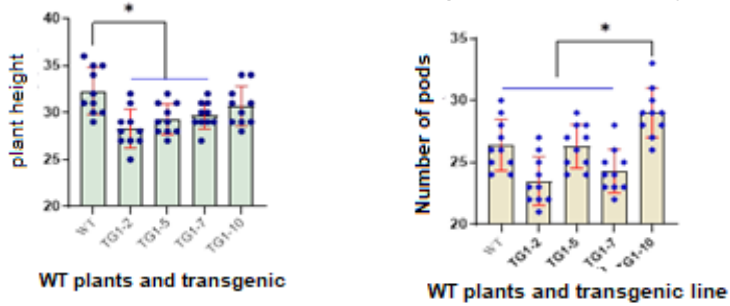
(A) (B)

Figure 3.22. Docking model of the AP2 domain and its hydrogen bonding interactions with the target DNA (GmP5CS promoter).

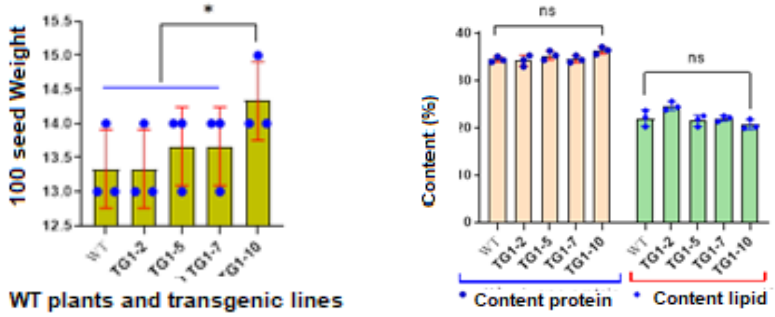
Hydrogen bonds (H-bonds) are critical for the affinity and specificity of the AP2-DNA interaction. As shown in Figure 3.22, the DNA-binding residues of the AP2 domain, including ARG2, GLY3, ARG5, ARG7, TRP9, LYS11, GLU15, ARG17, ARG24, TRP26, and THR29, form H-bonds with the promoter region containing *cis*-acting elements at DRE/CRT and GT-1 motifs including DG10, DC11, DC12, DG13, DC14, DT16, DG17. These findings confirm the binding of the DREB7 AP2 domain to the GmP5CS promoter.

3.3.4. Evaluation of promising transgenic soybean lines

The results of morphological, biochemical, and physiological analyses of the *GmDREB7* transgenic soybean lines (Figure 3.23) showed that, overall, the transgenic lines exhibited no significant differences in growth characteristics or seed quality compared with the WT plants. These findings indicate that the introduction of the transgene did not adversely affect the fundamental agronomic traits of soybean.



A **B**



C **D**

Figure 3.23. Comparison of selected morphological and biochemical characteristics of transgenic soybean lines and wild-type (WT) plants. **(A)** Plant height (cm), n = 10; **(B)** Number of pods per plant, n = 10; **(C)** 100-seed weight (g), n = 3; **(D)** Protein (% dry weight) and lipid (% dry weight) contents, n = 3.

Morphologically, plant height in the transgenic lines ranged from 28.20 to 30.70 cm, which was slightly lower than that of the WT plants (32.30 cm). Regarding the number of filled pods per plant, no significant differences were observed among the TG1-2, TG1-5, and TG1-7 lines compared with the WT plants. In contrast, the TG1-10 line exhibited the highest value (29.00 pods per plant), which was significantly greater than that of the WT plants (26.40 pods per plant; $p < 0.05$). For seed yield-related traits, the 100-seed weight of the transgenic lines ranged from 14.33 to 15.66 g and did not differ significantly from that of the WT plants.

Regarding the biochemical composition of the seeds, the analysis revealed that the protein and lipid contents of most *GmDREB7* transgenic soybean lines did not differ significantly from those of the WT plants ($p > 0.05$), indicating that seed quality remained stable following genetic transformation. Notably, the TG1-10 line exhibited the highest protein content (36.28%), exceeding that of the WT plants (34.38%).

More pronounced differences between the transgenic lines and WT plants were observed in physiological traits associated with stress responses. Specifically, proline accumulation was consistently higher in all transgenic lines under both drought and salt stress conditions, with particularly marked increases in the TG1-5 and TG1-10 lines. After 9 days of drought treatment, the proline content in TG1-10 reached 5.87

μmol g^{-1}, more than twice that of the WT plants (2.76 μmol g^{-1}). A similar trend was observed under salt stress, with the highest proline levels recorded in TG1-5 (7.37 μmol g^{-1}) and TG1-10 (6.88 μmol g^{-1}).

The combined evidence from gene expression, physiological, biochemical, and molecular docking analyses suggests that *GmDREB7* functions as a positive transcriptional regulator of the proline biosynthesis pathway through the activation of *GmP5CS*, thereby contributing to the maintenance of cellular water balance and enhancing drought and salt tolerance in soybean. The strong agreement between the experimental data and *in silico* analyses further highlights the considerable potential of the TG1-5 and TG1-10 lines as promising candidates for soybean improvement programs based on genetic transformation or molecular breeding approaches. Nevertheless, to conclusively demonstrate the direct regulatory mechanism of *GmDREB7*, additional molecular validation experiments, such as ChIP-qPCR and electrophoretic mobility shift assays (EMSA), should be conducted in future studies.

CONCLUSIONS AND RECOMMENDATIONS

1. Conclusions

1.1. The evolutionary relationships of the AP2 gene family, the DREB subfamily and the AP2 domain in *A. thaliana* and soybean were revealed through phylogenetic analysis, showing their distribution among monophyletic and paraphyletic groups. A total of 18 DREB genes were identified on 17 soybean chromosomes, whereas 12 DREB genes were located on 5 chromosomes of *A. thaliana*. The AP2 domains of DREB proteins contained 11 amino acid residues involved in promoter binding and were characterized by 12 conserved motifs, with RGRRWKERRWT and RGRRSKERRWT being the most prevalent. *GmDREB5* and *GmDREB7*, representing two distinct evolutionary clades, were selected for further analysis of their regulatory roles in target genes involved in the salt stress response.

1.2. A synthetic gene construct and an expression vector carrying the DREB5 encoding gene, *pFGC_GmDREB5*, belonging to the AP2 transcription factor family of soybean, were successfully generated. The construct harboring *GmDREB5* was successfully introduced into tobacco plants. *In silico* molecular docking analysis indicated that the

AP2 domain of DREB5 can bind to cis-regulatory elements in the promoter regions of target genes. Under salt stress conditions, *GmDREB5* expression enhanced the transcription of *NtP5CS* and *NtSODFe*, resulting in increased proline accumulation in transgenic tobacco plants.

1.3. The *pBI121_DREB7* construct containing the CaMV35S promoter and *GmDREB7* gene was successfully introduced into the soybean cultivar DT26 via *A. tumefaciens*-mediated transformation. Expression of *GmDREB7* enhanced the transcription of *GmP5CS* and promoted proline accumulation in transgenic soybean plants under salt stress conditions. Four transgenic soybean lines carrying *GmDREB7* (TG1-2, TG1-5, TG1-7, and TG1-10) exhibited significantly greater tolerance to both salt and drought stresses than wild-type (WT) plants, with TG1-5 and TG1-10 showing the highest levels of stress tolerance.

2. Recommendations

2.1. Future studies should focus on elucidating the contribution of *GmDREB5* to the enhancement of drought and salt stress tolerance in soybean.

2.2. Further evaluation of the two transgenic soybean lines, TG1-5 and TG1-10, should be conducted in subsequent generations to identify lines with superior stress tolerance for use as genetic resources in soybean breeding programs.

2.3. The *pFGC_GmDREB5* and *pBI121_DREB7* vectors have potential for application in genetic improvement programs targeting enhanced drought and salt tolerance in soybean as well as other legume species.