HOSTED BY

Contents lists available at ScienceDirect

Saudi Journal of Biological Sciences

journal homepage: www.sciencedirect.com

Original article

Insights into cucumber (*Cucumis sativus*) genetics: Genome-wide discovery and computational analysis of the Calreticulin Domain-Encoding gene (CDEG) family

Pravej Alam *, Thamer Albalawi

Department of Biology, College of Science and Humanities, Prince Sattam bin Abdulaziz University, Alkharj 11942, Saudi Arabia

ARTICLE INFO

Keywords:
Calreticulin domain encoding genes (CDEGs)
Unique genes
Cucumis sativus
Cold stress
Insilico analyses

ABSTRACT

Cucumber is an essential vegetable crop throughout the world. Cucumber development is vital for accomplishing both quality and productivity requirements. Meanwhile, numerous factors have resulted in substantial cucumber losses. However, the calreticulin domain-encoding genes (CDEGs) in cucumber were not well-characterized and had little function. In the genome-wide association study (GWAS), we recognized and characterized the CDEGs in *Cucumis sativus* (cucumber). Through a comprehensive study of *C. sativus*, our research has unveiled the presence of three unique genes, denoted as *CsCRTb, CsCRT3,* and *CsCNX1*, unevenly distributed on three chromosomes in the genome of *C. sativus*. In accordance to the phylogenetic investigation, these genes may be categorized into three subfamilies. Based on the resemblance with AtCDE genes, we reorganized the all *CsCDE* genes in accordance with international nomenclature. The expression analysis and *cis*-acting components revealed that each of *CsCDE* gene promoter region enclosed number of *cis*-elements connected with hormone and stress response. According to subcellular localization studies demonstrated that, they were found in deferent locations of the cell such as endoplasmic reticulum, plasma membrane, golgi apparatus, and vacuole, according to subcellular localization studies. Chromosomal distribution analysis and synteny analysis demonstrated the probability of segmental or tandem duplications within the cucumber CDEG gene family. Additionally, miRNAs displayed diverse modes of action, including mRNA cleavage and translational inhibition. We used the RNA seq data to analyze the expression of CDEG genes in response to cold stress and also improved cold tolerance, which was brought on by treating cucumber plants to an exogenous chitosan oligosaccharide spray. Our investigation revealed that these genes responded to this stress in a variety of ways, demonstrating that they may adapt quickly to environmental changes in cucumber plants. This study provides a base for further understanding in reference to CDE gene family and reveals that genes play significant functions in cucumber stress responses.

1. Introduction

The botanical family associated with flowering plants within the Cucurbitales order is recognized as the *Cucurbitaceae* family, commonly referred to as the gourd family. Cucumbers, gourds, melons, squashes, and pumpkins are among the 98 genera and over 975 species that make up this family (Chinnadurai, 2024). In the Cucurbitaceae family, cucumber is a monoecious annual crop that has been domesticated for more than 3,000 years (Ibrahim et al., 2023). The majority of species are tropical or temperate annual or perennial herbs that are sensitive to temperatures close to freezing, which restricts their range of cultivation and geographic distribution (Zhang et al., 2023). Due to its excellent yields and economic value, cucumber is frequently produced year-round in greenhouses in China. 87,805,086 tonnes were produced globally in 2019 on 2,231,402 ha of cultivable land, according to FAOSTAT. It is listed as the 10th most significant vegetable crop in the world. In 2019, China contributed 70,338,971 tonnes (80.11%) of the world's production from 1,258,370 ha (56.39%) of cultivable land (Sallam et al., 2021).

The endoplasmic reticulum (ER) is a cellular organelle that plays a vital role in numerous cellular processes, including protein synthesis, folding, and post-translational modification, as well as lipid metabolism. This organelle is involved in the storage, synthesis, and release of calcium ions, and its proper functioning is essential for normal cellular activity (Ali and Najeeb, 2023). The ER contains several proteins that are

* Corresponding author.
E-mail addresses: ps.alam@psau.edu.sa (P. Alam), t.albalawi@psau.edu.sa (T. Albalawi).

https://doi.org/10.1016/j.sjbs.2024.103959
Received 18 January 2024; Received in revised form 6 February 2024; Accepted 10 February 2024
Available online 13 February 2024

involved in these functions, such as calreticulin and calnexin, which act as molecular chaperones and are involved in the folding of newly synthesized glycoproteins. These proteins also play a role in the quality control system of the ER. The lectin site in both calreticulin and calnexin aids their association with newly synthesized glycoproteins (Crofts and Denecke, 1998; Danilczyk et al., 2000; Coe and Michalak, 2009). Calreticulin is strongly expressed in plants, particularly in tobacco, barley, Nicotiana plumbaginifolia, and maize embryogenesis, as well as in flower tissues such as sperm cells, pollen tubes, and anthers. It has been demonstrated that calreticulin in plant cells contributes to regeneration, pollen-pistil contact, and cell-to-cell transfer via the plasmodesmata (Joshi et al., 2019). In addition, calreticulin acts as a chaperone and is the primary Ca^{2+}-binding chaperone in the ER, responsible for the quality control of newly formed proteins and glycoproteins. Calreticulin-deficient and transgenic mice have shown that calreticulin and the ER can serve as upstream regulators in Ca^{2+}-dependent pathways that regulate cell differentiation and organ development (Michalak et al., 2009). Calreticulin and calnexin are involved in calcium homeostasis in plants. Calreticulin and calnexin are also involved in calcium homeostasis in plants, and calreticulin, as a calcium-binding protein, is responsible for the storage and release of calcium ions in the ER, which is important for various cellular processes (Trono and Pecchioni, 2022). Certain studies have also discovered CNX gene variants linked to variations in growth rate, tolerance to drought, or susceptibility to disease in various plant populations (Kang et al., 2023). CNX genes can have a variety of activities in addition to diverse expression patterns. For instance, certain CNX genes might control cell division and differentiation, whereas other CNX genes might control growth and water transport (Montpetit et al., 2023). Additionally, connexins, a class of proteins produced by the CNX gene, have been shown to be important for phagocytosis in plants. Gene knockout of connexins in *Arabidopsis thaliana* resulted in a significant reduction in phagocytosis (Marshall et al., 2015). Connexins play a role in the formation of gap junctions, which allow for the passage of ions, tiny molecules, and signaling molecules between adjacent cells. The varied features of the connexins that they encode, such as the size and composition of the gap junction channels, are probably the cause of the different functions of CNX genes (Liu et al., 2023). Calnexin and calreticulin have been identified and characterized in multiple plant species, including *Arabidopsis* (Huang et al., 1993), soybean (Goode et al., 1995), maize (Kwiatkowski et al., 1995), *Pisum sativum* (Ehtesham et al., 1999), and *Oryza sativa* (Sarwat and Naqvi, 2013). Maize, a crop plant, contains more than 60 CNX genes compared to the model plant *A. thaliana's* 26 (Sekhwal et al., 2015).

2. Material and method

2.1. Database search and sequence retrieval

Calreticulin domain (PF00262) of genes, which was used as a query sequence when looking for potential calreticulin protein sequences against *Cucumis sativus* in the genome database at PF00262 Phytozome v13 (https://phytozome-next.jgi.doe.gov) obatined from National Centre for Biotechnology Information (https://www.ncbi.nlm.nih.gov). The resulting amino acid (AA) sequences were confirmed employing NCBI CDD's default parameters (Conserved Domain Database) (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi).

2.2. Determination of physio-chemical properties

Sequences of Calreticulin domain-encoding genes (CDEGs) *C. sativus* and *A. thaliana* species were retrieved from Phytozome v13 (https://phytozome-next.jgi.doe.gov) database. The motif, gene position (start and endpoint), number of exons, genome and transcriptional length, molecular weight (Mw) and isoelectric point (pI) of sequences of Calreticulin domain-encoding genes (CDEGs) in *C. sativus* and *A. thaliana* species was confriemed by motif search tool (https://www.genome.jp/tools/motif/), ProtPram (https://web.expasy.org/protparam/) and plant ensemble (https://plants.ensembl.org/) online tools.

2.3. Comparative phylogenetic study

PF00262 (calreticulin) was utilized as the keyword to explore the evolutionary relationship among CNX and CRT members of the Cucurbitaceae family of four species (*C. sativus*, *C. maxima*, *C. pepo*, and *C. moschata*) together with Arabidopsis thaliana. Four Cucumis species CNX family members' protein sequences were provided into the MEGA-11 tool using the default settings in order to create a phylogenetic tree using the maximum likelihood technique and 1000 bootstrap replicates. To improve its visual appearance, it additionally employed an internet web server (https://itol.embl.de/).

2.4. Conserved domain motif analysis

The MEME (Multiple Expectation Maximization for Motif Elicitation) programs, available online at (http://meme.sdsc.edu/meme/website/intro.html), use the default set values for the best matching length value.

2.5. Cis-regulatory analysis

To extract *cis*-regulatory elements, the promotor regions from phytozome v13 (https://phytozome-next.jgi.doe.gov) were obtained using the online tool Plant Care (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). With a range of 5 to 20 bp, possible *cis*-elements were found and TB-tools were used to create a heat map to display the results.

2.6. Subcellular localization analysis

The WoLF PSORT (https://wolfpsort.hgc.jp/) database was used to determine the localized genes within cell of the *C. sativus*. Protein sequences are used as input in this database to provide a list of possible sites along with prediction scores. Finding the proteins' most likely position within the cell was the main goal of this work.

2.7. Exon-intron analysis

Calreticulin domain encoding genes (CDEGs) family genomic and CDS sequences were scanned for allocation of exons and introns by Gene Structure Display Server (GSDS) at (http://gsds.cbi.pku.edu.cn/).

2.8. MiRNA analysis

The Calreticulin domain-encoding genes family's target site was determined using the PmiREN database (https://www.pmiren.com/). The web server program psRNATarget (https://www.zhaolab.org/psRNATarget/) was used to evaluate the gene sequences of CDS along with older miRNA sequences, with the default parameters. The Cytoscape tool was used to envisage the associations among the goal genes and projected miRNA.

2.9. Syntney and dual synteny analysis

To investigate the syntenic relationships between paralogous genes in *C. sativus* and their orthologous counterparts in *Arabidopsis*, the researchers conducted an analysis of duplicate gene events using MCScanX. This analysis allowed them to identify duplicated genes within the genome. They then utilized the Advanced-Circos software integrated into TB-tools to create syntenic and dual syntenic maps, which helped them visually represent the connections and similarities between genes in these two plant species (Chen et al., 2020; Li et al., 2020).

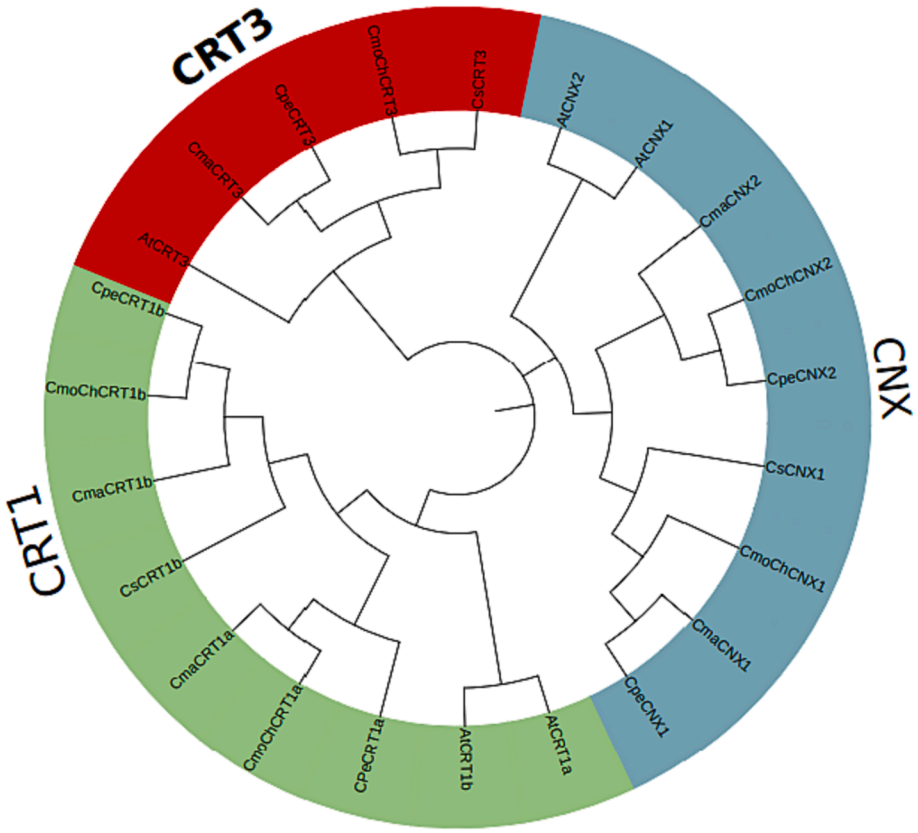


Fig. 1. Evolutionary analysis using MEGA-11 software investigated the phylogenetic relationships within the CDE gene family of *C. sativus, A. thaliana, C. maxima, C. pepo*, and *C. moschata*. The study employed iTol v6 to generate an interactive phylogenetic tree, facilitating a concise visualization of evolutionary patterns among the species' CDE genes.

2.10. Ka/Ks analysis

Based on the information available on Phytozome, we identified calreticulin domain-encoding genes (CDEGs) that were chromosomally located and duplicated using the TB-tools software. They then used the Ka/Ks calculator in TB-tools to determine the identical substitution rate (Ks), non-identical substitution rate (Ka), and Ka/Ks ratio within the repeated pairs of gene. The evolutionary divergence was calculated using the T = Ks/2 method, where T represents evolutionary divergence and Ks represents the synonymous substitution rate divided by 2 and multiplied by $1.5*10^{-8}$ (Han et al., 2021).

2.11. Transcriptomic analysis

C. sativus plants of three different genotypes (*CsCRTb*, *CsCRT3*, and *CsCNX1*) were chosen for this study's analysis to properly investigate how they responded to cold stress. The datasets with accession numbers GSE224757 and GSE210703 were obtained from this database from the NCBI GEO (Gene Expression Omnibus) database at (https://www.ncbi.nlm.nih.gov/geo/). The primary aim of this investigation was to elucidate the cold stress expression profiles of the CDEG family within these selected genotypes of *C. sativus*.

3. Results

3.1. Determination of physio-chemical properties

The investigation, as detailed in Table 1S, scrutinizes the genomic and transcriptional attributes of three genes: *CsCRT1b*, *CsCRT3*, and *CsCNX1*. Positioned on chromosome 2 in the forward direction, *CsCRT1b* is accompanied by *CsCRT3* on chromosome 4 in the reverse direction, and *CsCNX1* on chromosome 1 in the forward direction. The diverse lengths of both mRNA and peptides observed underscore the intricate protein complexities inherent in these genetic entities. *CsCRT1b* has the shortest coding sequence, *CsCRT3* is slightly longer, and *CsCNX1* has the longest. Isoelectric points (pI) indicate pH at which proteins carry no charge, with *CsCRT1b* at 4.45, *CsCRT3* at 5.8, and *CsCNX1* at 4.78. Molecular weights differ, influencing protein behavior in cellular processes. *CsCRT1b* has the lowest (48,386.42 KDa), *CsCNX1* the highest (60,732.99 KDa), and *CsCRT3* is intermediate (50,705.73 KDa). These findings deepen our understanding of these genes' genomic and biochemical characteristics, offering insights into their potential biological functions.

3.2. Comparative phylogenetic analysis

In the phylogenetic tree analysis, we investigated three distinct clades, namely CNX, CRT1, and CRT3. Among these clades, CNX and CRT3 exhibited a notably higher number of genes. Within the CNX clade, we identified the presence of the following genes: *AtCNX1, AtCNX2, CmaCNX2, CmoChCNX2, CpeCNX2, CsCNX1, CmaCNX1, CmoChCNX1,* and *CpeCNX1*. In the CRT1 group, we observed the presence of the following genes: *AtCRT1a, AtCRT1b, CmaCRT1a, CmoChCRT1a, CpeCRT1a, CsCRT1b, CmaCRT1b, CmoChCRT1b,* and *CpeCRT1b*. Lastly, within the CRT3 group, we identified the presence of the following genes: *AtCRT3, CmaCRT3, CmoChCRT3, CpeCRT3,* and *CsCRT3* Fig. 1.

3.3. Conserved domain analysis

All the discovered genes that code for the calreticulin domain have been found to contain the calreticulin superfamily and the calreticulin

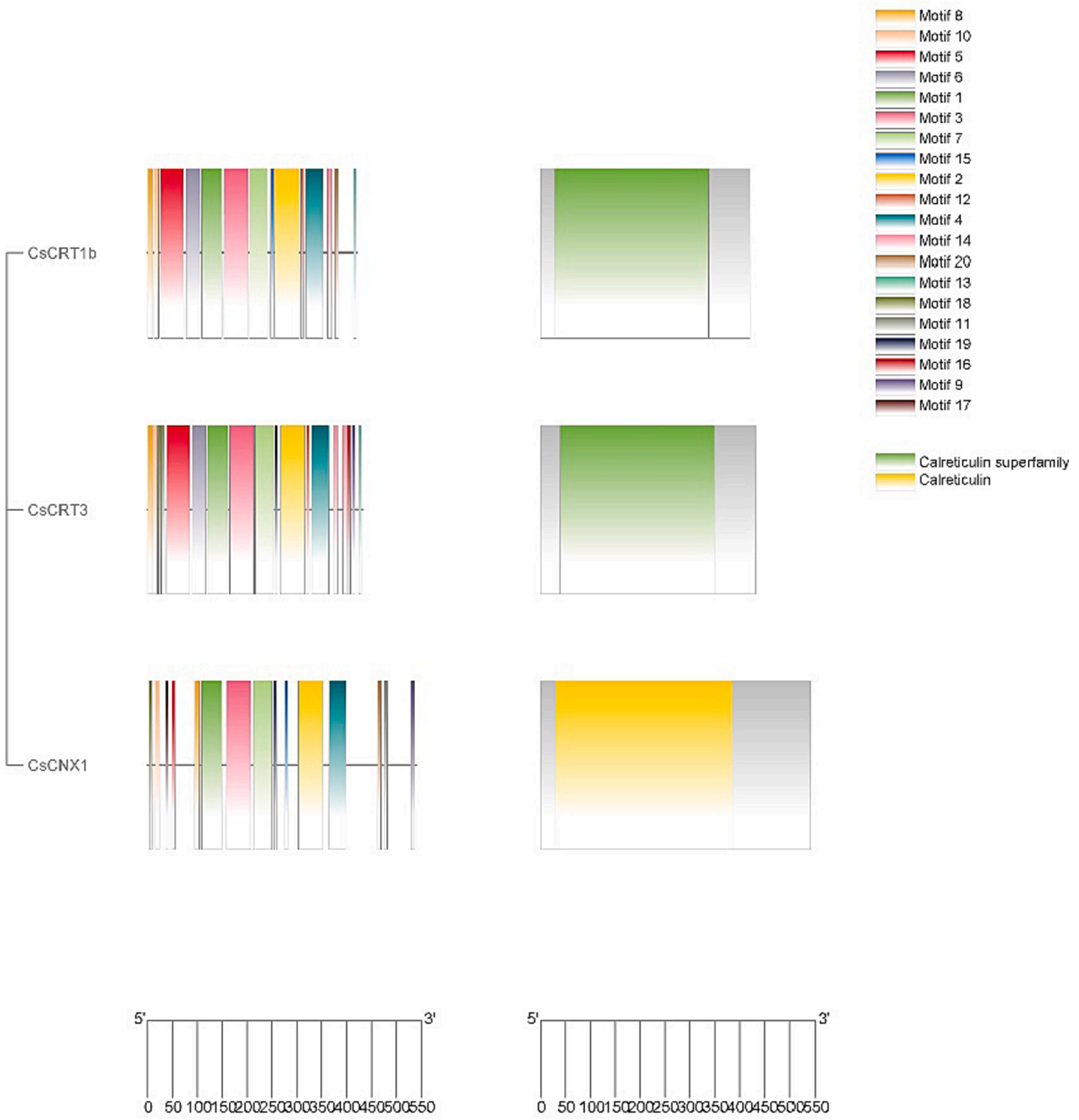


Fig. 2. A color-coded bar graph, generated using MEME version 5.5.2, illustrates the motif distribution analysis of *CsCDE* genes. Integration with a phylogenetic tree enhances comprehension of the evolutionary patterns and functional relationships among *CsCDE* genes, offering valuable insights.

domains. Notably, the CDEGs have been found to consistently include certain motifs, such as motif 1, motif 2, motif 3, motif 4, motif 7, and motif 15. Further investigations show that the remaining gene family members motif variation falls between the range of 5 and 20. The calreticulin superfamily domain of the *CsCRT1b* and *CsCRT3* genes is shown in green color in the Fig. 2, whereas the calreticulin domain of the *CsCNX1* gene is shown in yellow color. This important discovery clearly shows that the members of this gene family share a structural characteristic.

3.4. Cis-regulatory elements analysis

The results of the *cis*-regulatory component study showed (Fig. 3) that, in total, 36 *cis*-elements were identified. The CAAT-box *cis*-element had a high expression level, and its function was a general *cis*-acting element in both the promoter and enhancer regions. The TATA-box also exhibited a significant expression level, and it played a role as a core promoter element located around −30 from the transcription start site. Other *cis*-elements, such as GA-motif, ABRE, G-box, and MYB-like sequences, were present but did not show significantly higher expression levels. The functions of all the *cis*-elements mentioned are summarized in Table 2S (Supplementary Data).

3.5. Subcellular localization of CsCDE genes

CsCRT1b and *CsCRT3* exhibit high expression levels within the endoplasmic reticulum, while their expression is notably lower in the nucleus, cytoplasm, Golgi apparatus, vacuole, and peroxisomes. In

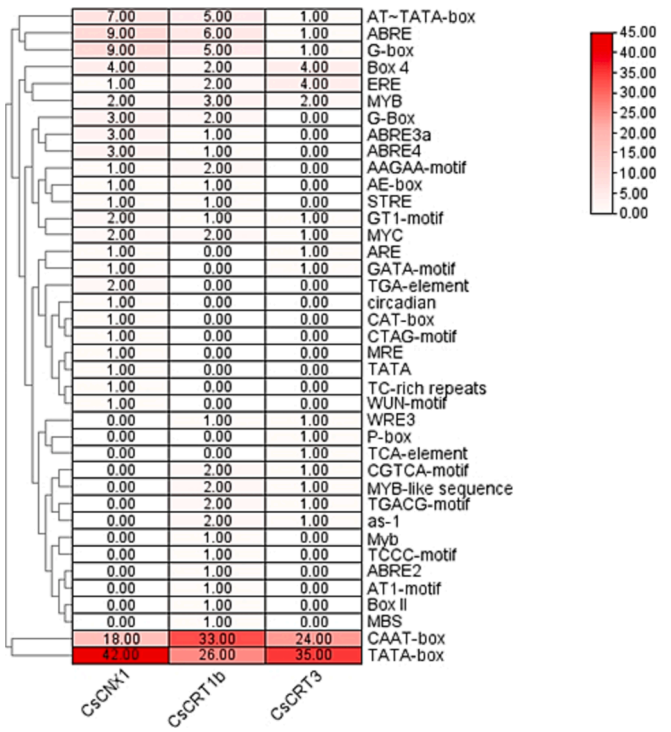


Fig. 3. Cucumber *CsCDE* gene promoters were analyzed, revealing correlations with various plant developmental processes. Statistical assessments quantified *cis*-regulatory element presence in each *CsCDE* gene.

contrast, the expression pattern of *CsCNX1* varies across different cellular compartments, displaying distinct levels of expression in the nucleus, cytoplasm, Golgi-apparatus, vacuole, and peroxisomes shown in Fig. 4.

3.6. Gene structure and mapping

The exons and introns structure provide the base for genes and aids in the study of evolutionary associations among the organisms and genes. A gene family can be recognized by their figures and dispersion patterns through evolutionary phylogenetic analysis, and a thorough demonstration of the exon–intron architecture of the CDEG family proteins showed that the gene structural pattern agreed with the phylogenetic analysis. CDEGs have between 6 (*CsCNX1*) and 13 (*CsCRT3*) exons. Exons and introns are approximately similar in all protein sequences Fig. 5.

The gene mapping analysis revealed a close linkage between these genes. Specifically, *CsCNX1* was mapped to chromosome 1, *CsCRT3* to chromosome 4, and *CsCRT1b* to chromosome 2. It suggests a potential functional similarity or association between these genes Fig. 6.

3.7. Syntney and dual synteny analysis

Our genome-wide synteny research produced persuasive results regarding *CsCRT1b*, *CsCRT3*, and *CsCNX1* conservation across various genomes. Three genes showed a high degree of syntenic preservation, with conserved gene order present in the genomes. Comparing genomes may be an efficient and quick technique to transfer genetic data from a taxon that has been extensively investigated to one that has not. In the CDEG family of proteins, three pairs of paralogous genes found through gene duplication may reveal important details about the growth of a gene family. Tandem duplications were not found, but segmental duplications were seen, which are principally responsible for the growth of plant gene families. A single chromosome contains adjacent homologous genes with no more than two intervening genes. *CsCRT3/CsCNX1* were detected in proximity on chromosome 4. *CsCRT1b/CsCRT3* were on chromosome 1 and *CsCNX1/CsCRT3* on chromosome 2. A comparative synteny diagram is shown in Fig. 7a. This finding signifies that there are regions on cucumber's chromosome 1 and *Arabidopsis thaliana's* Gm04 chromosome that contain orthologous genes. It shows that they share a common ancestry, suggesting similar functions and evolutionary origin Fig. 7b.

3.8. Ka and Ks analysis

The Ka/Ks ratio of 0.241 implies that the *CsCRT1b* and *CsCNX1* genes have been subjected to natural selection, highlighting the significance of maintaining functional integrity in these genes throughout their evolutionary history. Additionally, the data includes Ka and Ks values that provide information on the rates of synonymous and amino acid-changing mutations that accumulated between the two gene sequences during their evolutionary divergence. The estimated period since the divergence of the two sequences is also represented by the MYA number, which is 0.1656 Fig. 8.

3.9. MiRNA analysis

In this study, a total of 7 miRNAs were examined, all with a length of 22 base pairs (bp). These miRNAs are Csa-miRN811, Csa-miRN811, Csa-miR168, Csa-miRN967a, Csa-miRN967b, Csa-miRN967a, and Csa-miRN967b. Among these miRNAs, it was observed that Csa-miRN811 and Csa-miR168 are involved in suppressing translation. Furthermore, Csa-miRN967a, Csa-miRN967b, Csa-miRN967a, and Csa-miRN967b were found to be associated with the inhibition of cleavage. These results highlight the essential roles played by these miRNAs in post-transcriptional gene regulation, either by suppressing translation or by promoting the cleavage of their target mRNAs, as shown in Table 3S (Supplementary data).

3.10. Transcriptomic analysis

I. Gene expression profiling of cold-stressed cucumber plants

Notably, compared to conditions without cold exposure, there was a substantial up-regulation of *CsCRT1b* during a particular 3-hour cold treatment. The 3-hour cold treatment caused *CsCRT1b* to become overexpressed, which implies that *CsCRT1b* may be involved in the early response of cucumber seedlings to cold stress even when there is no external cold stress. This finding suggests that *CsCRT1b* may act in pathways unrelated to cold acclimatization through a constant or beginning level of expression. On the other hand, *CsCRT3*, a different cold-responsive gene, showed a unique pattern when the whole range of cold exposure treatments (3 h, 12 h, and 24 h) were examined. *CsCRT3* continuously showed up-regulation in comparison to situations without cold stress during these different treatment durations. When cold stress is present, *CsCRT3* is consistently raised, suggesting that it actively participates in the organism's response to cold stress and subsequent cold acclimatization. *CsCRT1b* and *CsCRT3's* differential regulation in response to cold exposure suggests the diverse roles that they may play in the intricate process of cold acclimatization. Even in the absence of external cold stress, *CsCRT1b* appears to have innate up-regulation, suggesting a possible function in processes unrelated to cold acclimatization. However, in response to cold stress, *CsCRT3* shows a distinct and steady up-regulation, highlighting its significance in the adaptive response to low temperatures. Further research is needed to completely understand how the differential regulation of CsCRT1b and CsCRT3 influences how cucumber seedlings respond to cold stress in general and the acclimatization processes that accompany it (Fig. 9a, b).

II. Gene expression profiling of improved cold tolerance in cucumber (*Cucumis sativus* L.) caused by exogenous chitosan oligosaccharide

The use of an exogenous chitosan oligosaccharide spray greatly increased the seedlings' resistance to cold stress, according to a gene

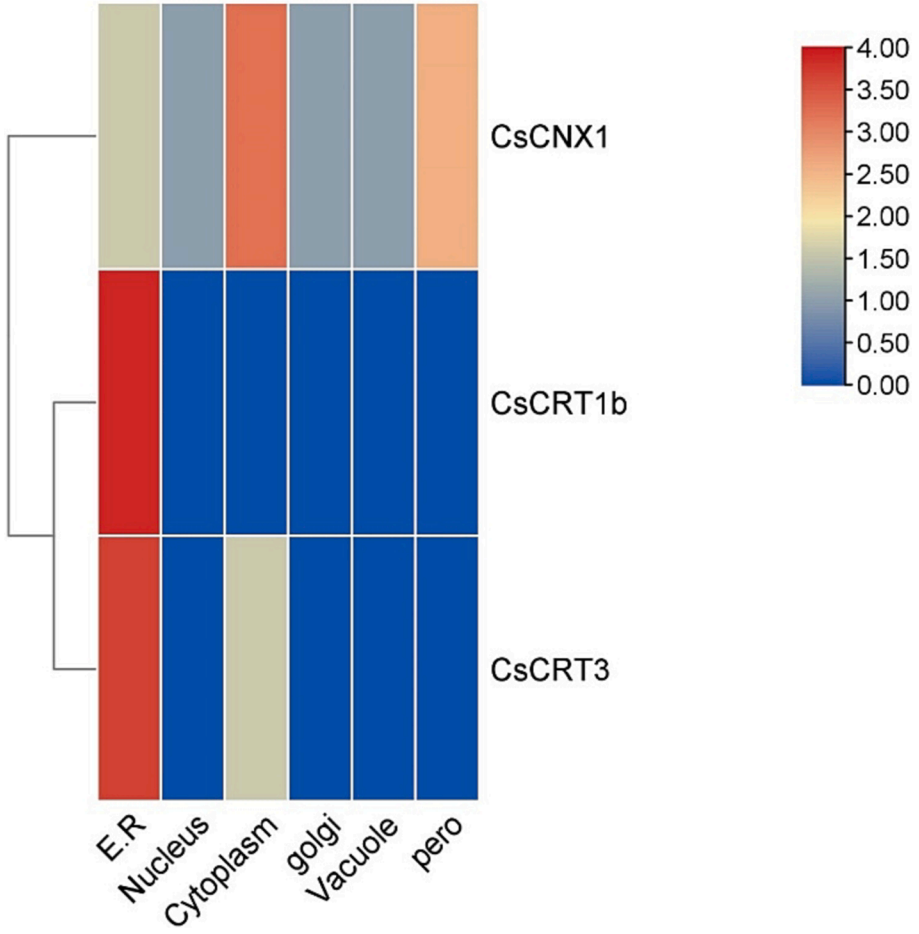


Fig. 4. The subcellular localization analysis of *CsCDE* genes reveals a predominant presence in the endoplasmic reticulum and cytoplasm, with a majority of proteins exhibiting a red color, indicating localization within these cellular compartments.

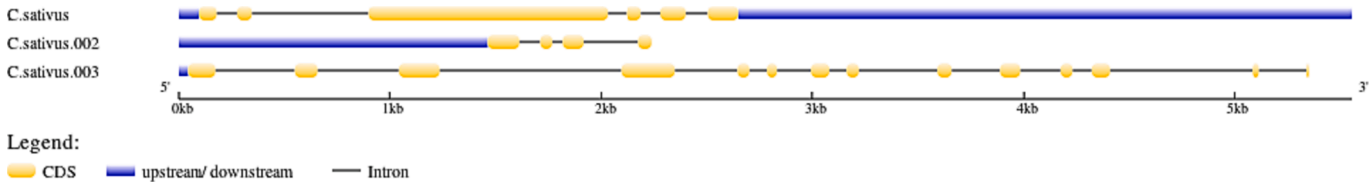


Fig. 5. The depiction of intron–exon structures utilized yellow and black shapes to denote exons within the genes.

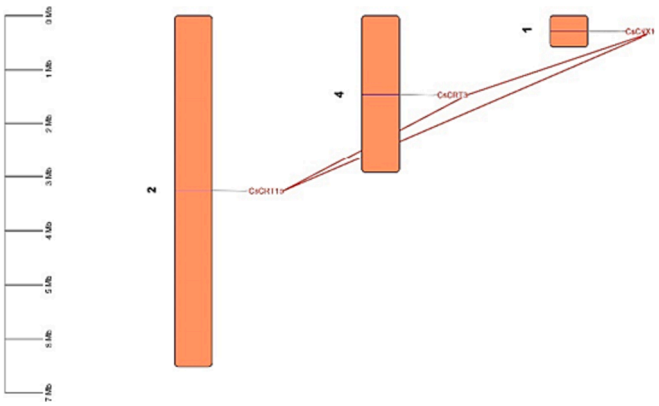


Fig. 6. The image depicts the *CcDEG* on different chromosomes and their linkage.

expression investigation of cucumbers (*C. sativus* L.). The research aimed to examine the molecular responses of cucumber seedlings to treatments of 3, 12, and 24 h of cold exposure in comparison to a control group that was not subjected to any form of cold stress. We examined how these different cold exposure scenarios affected the expression patterns of two specific cold-responsive genes, *CsCRT1b* and *CsCRT3*.

In particular, *CsCRT1b* showed a notable increase after an exogenous chitosan oligosaccharide spray application and a 12-hour cold treatment as compared to no cold exposure. The results of this analysis indicate that the use of an exogenous chitosan oligosaccharide spray treatment may have positively regulated *CsCRT1b*, hence improving the seedlings' resistance to cold stress. The observed overexpression of *CsCRT1b* in cucumber seedlings could suggest that this gene plays a function in reducing the negative consequences of cold stress and enhancing their ability to withstand it.

Additionally, *CsCRT3* consistently showed elevation when the gene expression profiles were compared to the control group without cold stress for all cold exposure treatments, including 3 h, 12 h, and 24 h. This

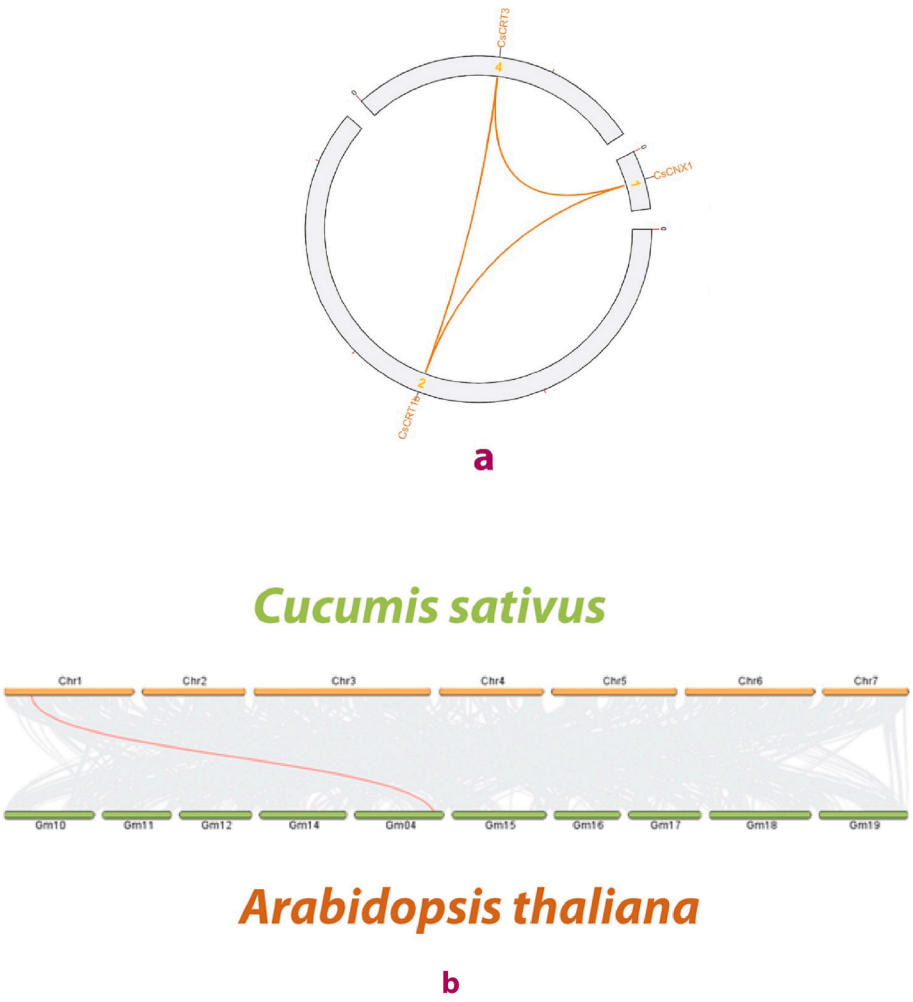


Fig. 7. **(a)** The arrangement of *CsCDE* genes across *C. sativus* chromosomes is depicted, with potential gene duplications represented by interchromosomal lines. **(b)** illustrates the chromosomal distribution and intra-chromosomal linkages of CDE genes, highlighting duplicate gene pairs with red lines and synteny blocks within the cucumber genome with gray lines.

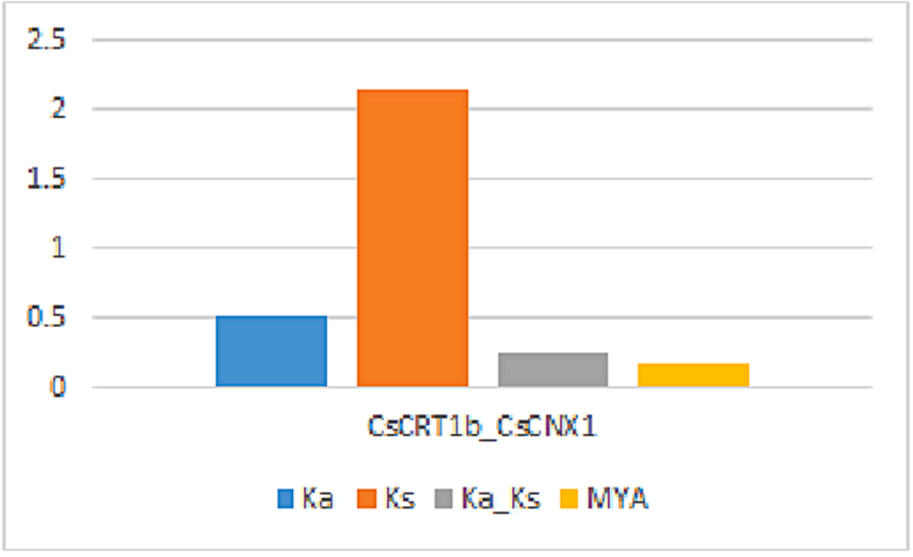
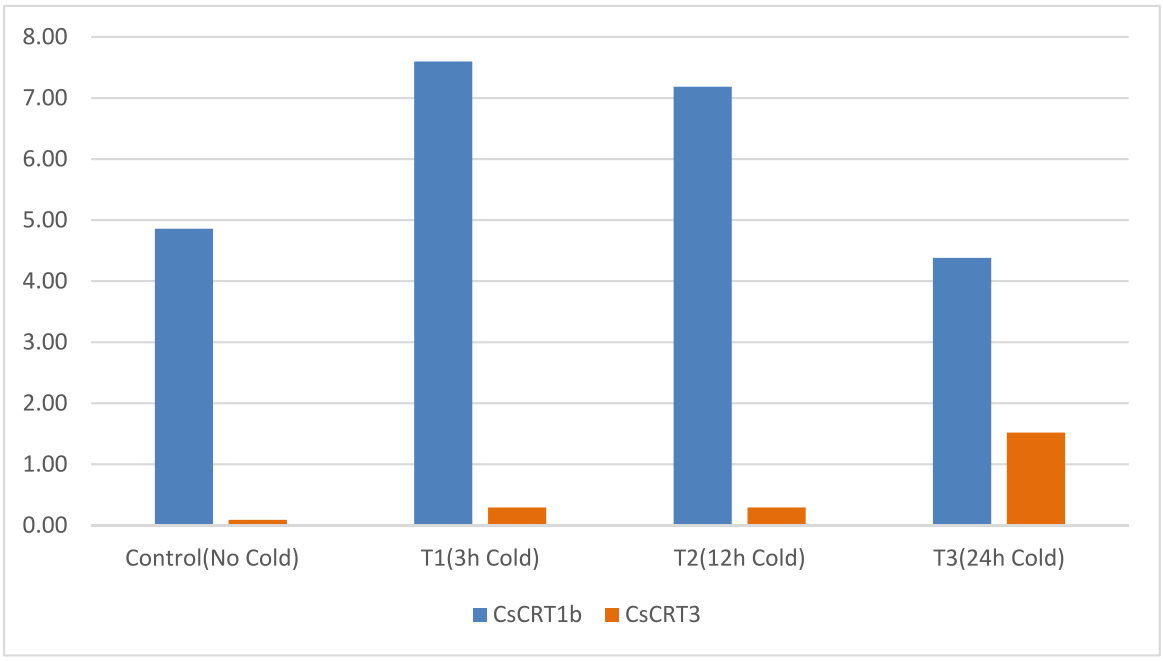
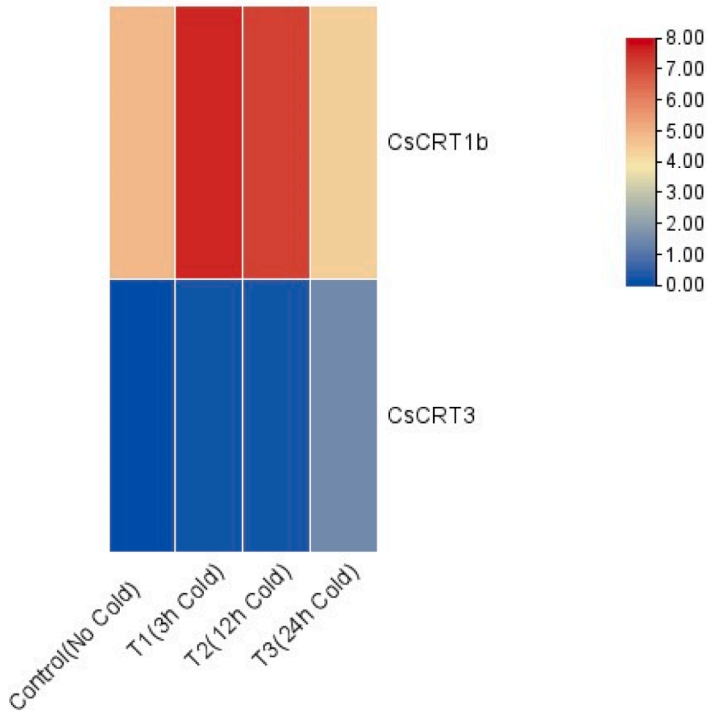


Fig. 8. The figure shows the non-synonymous substitution rate (ka) and synonymous substitution rate (ks), their ratio, and the evolutionary divergence period.



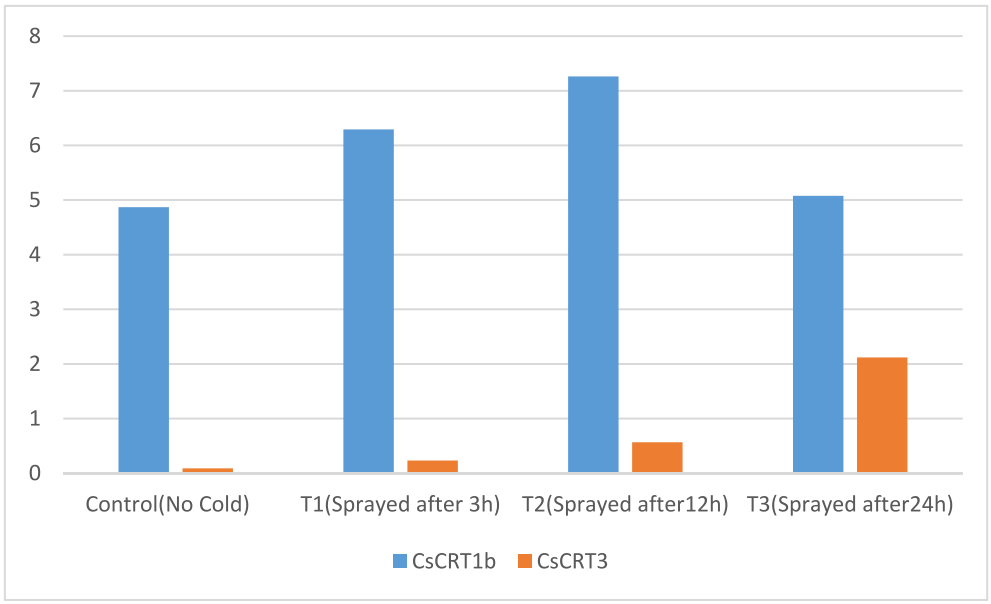
a



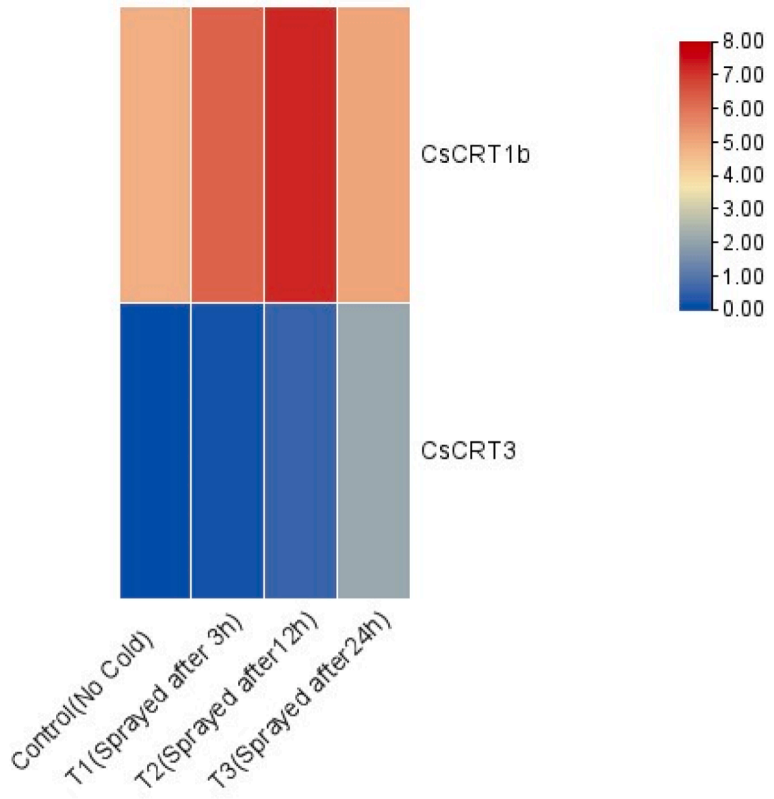
b

Fig. 9. (a & b): The effect of cold stress on *CsCRT1b* and *CsCRT3* is exhibited in the figure 9a and 9b, as is the response to non-cold and cold acclimatization.

steady elevation of *CsCRT3* over a range of cold exposure times suggests that it plays a part in the adaptation of cucumber seedlings to cold environments and shows that it is actively involved in the response to cold stress. Notably, *CsCRT3* showed a markedly elevated level of up-regulation following a 24-hour exposure to cold, underscoring its significance in longer-term cold stress situations. These findings highlight the positive impact of exogenous chitosan oligosaccharide spray on improving the cold stress tolerance of cucumber seedlings. While



a



b

Fig. 10. (a & b): The comprehensive response of *CsCRT1b* and *CsCRT3* is shown by the exposure of an exogenous chitosan oligosaccharide spray that increases resistance in seedlings in response to cold stress.

CsCRT3 consistently and significantly up-regulates itself over a wide range of cold exposure durations, the different expression patterns of *CsCRT1b* and *CsCRT3* indicate that these genes play different roles in the molecular response to cold stress.

4. Discussion

Cucumber is an important vegetable crop for several underdeveloped countries, but it is susceptible to yield reductions due to various environmental stressors, like biotic and abiotic (Parvathi et al., 2022). Abiotic stress in cucumber may trigger distinct transcriptomic

responses, resulting in significant yield loss. Transcription factors (TFs) participate an important role in controlling plant responses to various stressors, which helps to minimize such losses (Tian et al., 2022). However, the importance of the *CsCDE* gene family in cucumber is yet unclear. Although there are no specific studies on the role of *CsCDE* gene family in cucumber, there are studies on the effects of environmental stressors on cucumber yield and quality. Conventional methods in addition to molecular approaches as well as technologies employed by bioinformatics can enhance the fruit production and quality characteristics of *C. sativus*.

Cucumber's CDE gene family has been identified and investigated via a genome-wide analysis. To understand the relationship between cucumber CDEG members and other species, a phylogenetic tree of CDEG family members of *C. sativus, C. maxima, C. pepo, C. moschata and A. thaliana* was constructed, which divides the 23 CDEGs into three groups (CNX1,CRT3 and CRT1b; Fig.1 and 2). The three *CsCDEG* members fell into all three groups, which suggests that the *CsCDEG* structure and their function are highly conserved during plant evolution (Li et al., 2022). Subcellular localization investigation demonstrated that *CsCDE* genes have been identified in a wide range of organelles. Over half of the genes were found in the ER and cytoplasm, indicating that CsCDE proteins might have significant roles in these organelles (Mariani et al., 2003; Fig. 4). *Cis*-regulatory elements typically appear in the promoter region of genes and are required for transcriptional regulation of gene expression. The *cis*-regulatory elements like ABRE, MYC, TCA-element, GARE-motif and ERE are involved in phytohormones responses, mainly responds to abscisic acid, jasmonic acid, salicylic acid, gibberellin, and estrogen respectively (Maqsood et al., 2022). The binding sites GA-motif, box S, C, GARE-motif, chs-CMA1a, HD-Zip3, MBS, G-box, MYB-like sequence, Sp1, C, I-box, ATCT-motif, TCA-element, O2-site, E2Fb, and TATA-box are involved in various stresses like light stress, drought stress, and other abiotic and biotic stresses (Fig. 3). These binding sites have a role in regulating the function of *CsCDEG* under various stresses. Understanding the function of these binding sites is essential for developing strategies to improve plant growth and stress tolerance (Garg et al., 2015). Gene duplications drive genetic innovation by creating redundancy, preventing duplicated genes from selective pressure. This redundancy acts as a genetic 'spare part', ensuring continued functionality in the face of mutations affecting the original gene copy (Magadum et al., 2013). Synteny analysis is a valuable method for studying various gene duplication events in plants, including whole-genome duplication, tandem duplication, and proximal duplication. It involves assessing gene arrangement and sequencing across chromosomes or within a genome, helping researchers identify duplications and understand gene family evolution and functional diversity (Qiao et al., 2019; Fig. 5 and 6). Furthermore, the synteny blocks within Cucumber's genome were established and researched individually to clarify their links. Two genes on chromosomes 4, 2, and 1, and one gene on each of chromosomes 1, 2, and 4 were duplicated in cucumber (Zhang et al., 2023; Fig. 7).

The value of the Ka/Ks ratio for CsCRT1b and CsCNX1 is 0.241, which suggests that these genes have undergone natural selection to maintain their functional integrity throughout their evolutionary history. This is a crucial aspect of their evolution, as it ensures that they continue to function effectively. Research has also been conducted on wheat plants, and the Ka/Ks ratio for all duplicated SRO gene pairs was found to be less than 1 (Jiang et al., 2020). These gene pairs experienced stronger selective pressure, but they did not exhibit any significant differences in their functions during evolution. According to analysis of the dual synteny block between the genomes of *C. sativus* and *A. thaliana*, one gene was duplicated in *Arabidopsis*. MicroRNAs (miRNAs) are vital regulatory molecules in plants that take part in all biological developments, including plant growth, yield, and responses to stress. They are extremely conserved and display detailed functions (Zhang et al., 2022). A putative miRNA study showed that four miRNAs targeted *CsCDE* genes (Fig. 9). MiRNAs, such as miRN811, miRN967a, miRN967b, and miR168, play significant functions within plant biology, regulating developmental processes, stress responses, and universal gene expression. Their numerous functions emphasize how crucial they are in the growth of plants, stress tolerance, and adaptability to environmental changes (Millar, 2020). The significance of the CDE gene family group in cucumber has been demonstrated through an in depth transcriptome study using RNA sequencing data from the NCBI GEO database (Fig. 9 and 10). A detailed study revealed that CRT1b and CNX1 were highly expressed during cucumber's cold stress (Duan et al., 2022) and improved cold tolerance by using exogenous chitosan oligosaccharide spray (Tan et al., 2023). Not only does this remarkable finding pave the way for future research on enhancing the productivity and quality of cucumber, but it also provides a potential basis for developing new cucumber varieties by mutating CDEGs with improved characteristics. The precise molecular processes underlying the reactions of *CsCRT1b* and *CsCRT3* to cold stress and exogenous therapies need to be further explored.

5. Conclusion

This study comprised an in-depth investigation of the CsCDE gene family in the cucumber genome. To develop cold-resistant cultivars for agricultural use, this research advances our knowledge of how exogenous treatments might modify gene expression and improve cold stress tolerance in cucumber seedlings Fig. 10(a, b). In summary, a total of 23 CDE proteins were depicted in *C. sativus, C. maxima, C. pepo,* and *C. moschata*, together with *A. thaliana*. These genes were categorized into three subgroups named CNX, CRT3, and CRT1b by their functional and structural properties. In agreement with the occurrence of *CsDEGs* in the identical clade as *A. thaliana*, they are known to role in protein folding in the ER and are expressed in most tissues throughout development, alleviating ER stress and PAMP signaling, respectively. *CsCDEG* gene structure has been demonstrated to be primarily conserved throughout evolution. *Cis*-acting element analysis indicated that these genes might respond to multiple stresses such as cold stress, heat stress, pathogen stress, etc. *CsCRT1b* and *CsCNX1* were observed to have significant overexpression during cold stress. Future aspects of this investigation might involve generating *C. sativus* plants cold-tolerant by altering *CsCRT1b* and *CsCNX1* and modifying their gene expression patterns to over-expression in cold stress conditions. In summary, above research gives an excellent foundation for future research into the workings of *CsCDE* genes in cucumber. However, further study is required, such as gene cloning and useful analysis, to validate the importance of these genes in many physiological and biological processes.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2023/01/26500).

Funding

Prince Sattam bin Abdulaziz University project number (PSAU/2023/01/26500).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.

org/10.1016/j.sjbs.2024.103959.

References

Ali, S., Najeeb, M., 2023. Navigating the endoplasmic reticulum: new insights and emerging concepts. IntechOpen. 105737 https://doi.org/10.5772/intechopen.

Chen, C., Chen, H., Zhang, Y., Thomas, H.R., Frank, M.H., He, Y., Xia, R., 2020. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol. Plant 13, 1194–1202. https://doi.org/10.1016/j.molp.2020.06.009.

Chinnadurai C., 2024. Cucurbitaceous vegetables-I. In Viral Diseases of Field and Horticultural Crops, Elsevier pp. 411-417. https://doi.org/10.1016/b978-0-323-90899-3.00030-6.

Coe, H., Michalak, M., 2009. Calcium binding chaperones of the endoplasmic reticulum. Gen. Physiol. Biophys. 28, 96–103. https://doi.org/10.1016/S0143416002001884.

Crofts, A.J., Denecke, J., 1998. Calreticulin and calnexin in plants. Trends in Plant Sci. 3, 396–399. https://doi.org/10.1016/S1360-1385(98)01312-0.

Duan, X., Yu, X., Wang, Y., Fu, W., Cao, R., Yang, L., Ye, X., 2022. Genome-wide identification and expression analysis of glutathione S-transferase gene family to reveal their role in cold stress response in cucumber. Front. Gen. 13, 1009883. https://doi.org/10.3389/fgene.2022.1009883.

Ehtesham, N.Z., Phan, T.N., Gaikwad, A., Sopory, S.K., Tuteja, N., 1999. Calnexin from *Pisum sativum*: cloning of the cDNA and characterization of the encoded protein. DNA and Cell Biol. 18, 853–862. https://doi.org/10.1089/104454999314854.

Garg, G., Yadav, S., Ruchi, Y.G., 2015. Key roles of calreticulin and calnexin proteins in plant perception under stress conditions: a review. Adv. Life Sci. 5, 18–26. https://doi.org/10.5923/j.als.20150501.03.

Goode, J., Settlage, S., Wilson, R., Dewey, R., 1995. Isolation of a calnexin homolog from developing soybean seeds. Plant Physiol. 108, 1341. https://doi.org/10.3390/plants9030387.

Han, N., Tang, R., Chen, X., Xu, Z., Ren, Z., Wang, L., 2021. Genome-wide identification and characterization of WOX genes in *Cucumis sativus*. Genome. 64, 761–776. https://doi.org/10.1139/gen-2020-0029.

Huang, L., Franklin, A., Hoffman, N., 1993. Primary structure and characterization of an *Arabidopsis thaliana* calnexin-like protein. J. Biol. Chem. 268, 6560–6566. https://doi.org/10.1016/S0021-9258(18)53287-7.

Ibrahim, A.A., Salem, K.F., Abdein, M.A., Ramadan, S.M., 2023. Advances in Summer Squash (*Cucurbita pepo* L.) Molecular Breeding Strategies. In Smart Plant Breeding for Vegetable Crops in Post-genomics Era, Springer, pp. 163-215. https:// 10.1007/978-981-19-5367-5_8.

Jiang, W., Geng, Y., Liu, Y., Chen, S., Cao, S., Li, W., Chen, H., Ma, D., Yin, J., 2020. September). Genome-wide identification and characterization of SRO gene family in wheat: Molecular evolution and expression profiles during different stresses. Plant Physiol. Biochem. 154, 590–611. https://doi.org/10.1016/j.plaphy.2020.07.006.

Joshi, R., Paul, M., Kumar, A., Pandey, D., 2019. Role of calreticulin in biotic and abiotic stress signalling and tolerance mechanisms in plants. Gene. 714, 144004 https://doi.org/10.1016/j.gene.2019.144004.

Kang, Y., Tong, J., Liu, W., Jiang, Z., Pan, G., Ning, X., Yang, X., Zhong, M., 2023. Comprehensive analysis of major latex-like protein family genes in cucumber (Cucumis sativus L.) and their potential roles in phytophthora blight resistance. Int. J. Mol. Sci. 24, 784. https://doi.org/10.3390/ijms24010784.

Kwiatkowski, B., Zielińska-Kwiatkowska, A., Migdalski, A., Kleczkowski, L., Wasilewska, L., 1995. Cloning of two cDNAs encoding calnexin-like and calreticulin-like proteins from maize (Zea mays) leaves: identification of potential calcium-binding domains. Gene. 165, 219–222. https://doi.org/10.1016/0378-1119(95)00537-G.

Li, C., Dong, S., Liu, X., Bo, K., Miao, H., Beckles, D.M., Zhang, S., Gu, X., 2020. Genome-wide characterization of cucumber (Cucumis sativus L.) GRAS genes and their response to various abiotic stresses. Horticulturae. 6, 110. https://doi.org/10.3390/horticulturae6040110.

Li, J., Li, H., Quan, X., Shan, Q., Wang, W., Yin, N., Wang, S., Wang, Z., He, W., 2022. Comprehensive analysis of cucumber C-repeat/dehydration-responsive element binding factor family genes and their potential roles in cold tolerance of cucumber. BMC Plant Biol. 22, 1–13. https://doi.org/10.1186/s12870-022-03664-z.

Liu, Y., Cao, M., Yan, X., Cai, X., Li, Y., Li, C., Xue, T., 2023. Genome-wide identification of gap junction (connexins and pannexins) genes in black rockfish (Sebastes schlegelii): Evolution and immune response mechanism following challenge. Fish & Shellfish Imm. 132, 108492 https://doi.org/10.1016/j.fsi.2022.108492.

Magadum, S., Banerjee, U., Murugan, P., Gangapur, D., Ravikesavan, R., 2013. Gene duplication as a major force in evolution. J. Genet. 92, 155–161. https://doi.org/10.1007/s12041-013-0212-8.

Maqsood, H., Munir, F., Amir, R., Gul, A., 2022. Genome-wide identification, comprehensive characterization of transcription factors, cis-regulatory elements, protein homology, and protein interaction network of DREB gene family in *Solanum lycopersicum*. Front. Plant Sci. 13, 1031679. https://doi.org/10.3389/fpls.2022.1031679.

Mariani, P., Navazio, L., Zuppini, A., 2003. Calreticulin and the endoplasmic reticulum in plant cell biology. In Calreticulin: Second Edition, Springer, pp. 94-104. https://doi.org/10.1007/978-1-4419-9258-1_10.

Marshall, R.S., Li, F., Gemperline, D.C., Book, A.J., Vierstra, R.D., 2015. Autophagic degradation of the 26s proteasome is mediated by the dual atg8/ubiquitin receptor rpn10 in arabidopsis. Mol. Cell. 58, 1053–1066.

Michalak, M., Groenendyk, J., Szabo, E., Gold, L.I., Opas, M., 2009. Calreticulin, a multi-process calcium-buffering chaperone of the endoplasmic reticulum. Bioch. J. 417, 651–666. https://doi.org/10.1042/BJ20081847.

Millar, A.A., 2020. The function of miRNAs in plants. Plants 9, 198. https://doi.org/10.3390/plants9020198.

Montpetit, J., Clúa, J., Hsieh, Y.-F., Vogiatzaki, E., Müller, J., Abel, S., Strasser, R., Poirier, Y., 2023. Endoplasmic reticulum calnexins participate in the primary root growth response to phosphate deficiency. Plant Physiol. 191, 1719–1733. https://doi.org/10.1093/plphys/kiac595.

Parvathi, M., Antony, P.D., Kutty, M.S., 2022. Multiple stressors in vegetable production: insights for trait-based crop improvement in cucurbits. Front. Plant Sci. 13, 861637 https://doi.org/10.3389/fpls.2022.861637.

Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., Zhang, S., Paterson, A.H., 2019. Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. Genome Biol. 20, 1–23. https://doi.org/10.1186/s13059-019-1650-2.

Sarwat, M., Naqvi, A.R., 2013. Heterologous expression of rice calnexin (OsCNX) confers drought tolerance in Nicotiana tabacum. Mol. Biol. Rep. 40, 5451–5464. https://doi.org/10.1007/s11033-013-2643-y.

Sekhwal, M.K., Li, P., Lam, I., Wang, X., Cloutier, S., You, F.M., 2015. Disease resistance gene analogs (RGAs) in plants. Int. J. Mol. Sci. 16, 19248–19290. https://doi.org/10.3390/ijms160819248.

Tan, C., Li, N., Wang, Y., Yu, X., Yang, L., Cao, R., Ye, X., 2023. Integrated physiological and transcriptomic analyses revealed improved cold tolerance in cucumber (Cucumis sativus L.) by exogenous chitosan oligosaccharide. Int. J. Mol. Sci. 24, 6202. https://doi.org/10.3390/ijms24076202.

Tian, J., Li, Y., Hu, Y., Zhong, Q., Yin, J., Zhu, Y., 2022. Mining the roles of cucumber DUF966 genes in fruit development and stress response. Plants. 11, 2497. https://doi.org/10.3390/plants11192497.

Trono, D., Pecchioni, N., 2022. Candidate genes associated with abiotic stress response in plants as tools to engineer tolerance to drought, salinity and extreme temperatures in wheat: An Overview. Plants. 11, 3358. https://doi.org/10.3390/plants11233358.

Zhang, Y., Lin, D., Yan, R., Xu, Y., Xing, M., Liao, S., Wan, C., Chen, C., Zhu, L., Kai, W., 2023b. Amelioration of chilling injury by fucoidan in cold-stored cucumber via membrane lipid metabolism regulation. Foods. 12, 301. https://doi.org/10.3390/foods12020301.

Zhang, F., Yang, J., Zhang, N., Wu, J., Si, H., 2022. Roles of microRNAs in abiotic stress response and characteristics regulation of plant. Front. Plant Sci. 13, 919243 https://doi.org/10.3389/fpls.2022.919243.

Zhang, M., Zhang, W., Zheng, Z., Zhang, Z., Hua, B., Liu, J., Miao, M., 2023a. Genome-wide identification and expression analysis of NPF genes in cucumber (Cucumis sativus L.). Plants. 12, 1252. https://doi.org/10.3390/plants12061252.