

# GENOME WIDE IDENTIFICATION AND CHARACTERIZATION OF EDS1 TRANSCRIPTION FACTOR GENE FAMILY IN *CUCUMIS SATIVUS*

**MUHAMMAD MUNEEB HASHMI**

Department of Horticulture, Faculty of Agricultural Sciences, University of the Punjab, P.O BOX, Lahore, Pakistan. Email: muneebhashmi28@gmail.com

**MUHAMMAD SHAFIQ**

Department of Horticulture, Faculty of Agricultural Sciences, University of the Punjab, P.O BOX, Lahore, Pakistan. Email: shafiq.iags@pu.edu.pk

**MUHAMMAD ASHFAQ**

Department of Plant Breeding and Genetics, Faculty of Agricultural Sciences, University of the Punjab, P.O BOX, Lahore, Pakistan. Email: ashfaq.iags@pu.edu.pk

## Abstract

This study focused on identifying and analyzing the EDS1 genes in cucumber using various bioinformatics tools and databases. The EDS1 (the ENHANCED DISEASE SUSCEPTIBILITY 1) and PAD4 (PHYTOALEXIN DEFICIENT 4) proteins play a significant role in regulating gene expression in response to biotic and abiotic stresses. The study identified three CsEDS1 genes in cucumber through BLAST-p analysis and analyzed their gene structures, conserved motifs, and domains. The phylogenetic analysis was used to determine the evolutionary relationship between CsEDS1 and other related species. Cis-regulatory element analysis was performed to evaluate the putative function of various cis-elements in genes, and transcriptome analysis under different conditions showed the expression patterns of CsEDS1 genes. Subcellular localization analysis showed that all CsEDS1 sequences were localized in the nucleus, cytoplasm, and chloroplast. The study used Ks/Ka analysis to reveal that purification selection by the environment was the result of the evolution of CsEDS1 genes. Chromosomal mapping and synteny analysis showed segmental duplication of CsEDS1 genes. These results provide valuable insights into the evolution and function of CsEDS1 genes in cucumber, highlighting their significant role in plant survival and adaptation to different stress conditions. The findings of this study have implications for crop improvement and stress tolerance in plants. Understanding the molecular characteristics and evolutionary relationships of CsEDS1 genes can help develop strategies for improving plant resistance to biotic and abiotic stresses. The study also sheds light on the function of EDS1 and PAD4 proteins in regulating gene expression in response to environmental stresses, contributing to basal immunity through transcriptional regulation, and demonstrating different modes of action and pathway connectivities.

**Keywords:** Expression pattern, Genome wide, Potato, Transcription factor, *EDS* gene family, Transcription factors

## INTRODUCTION

Transcription factors (TFs) are regulatory proteins that play a crucial role in controlling gene expression in response to environmental changes. In plants, TFs are essential for coordinating the activation of genes involved in the defense mechanisms against biotic and abiotic stresses. TFs can act as activators or repressors of gene expression, binding to specific DNA sequences in the promoter regions of target genes to either enhance or suppress their transcription.

In response to biotic stresses, such as pathogen attacks, plants activate a series of signaling pathways that lead to the activation of TFs, which in turn induce the expression of genes encoding defense-related proteins, including pathogenesis-related (PR) proteins, protease inhibitors, and phytohormones. These proteins contribute to the plant's basal immunity against various pathogenic strains recognized by specific nucleotide-binding leucine-rich repeat (NLR) receptors, as well as effector-triggered immunity (ETI). TFs also play a crucial role in the modulation of jasmonic acid (JA), salicylic acid (SA), and ethylene (ET) signaling pathways, which are key components of the plant's immune response.

In addition to biotic stresses, TFs are also involved in the regulation of gene expression in response to abiotic stresses, including drought, salinity, and temperature stress. For instance, dehydration-responsive element-binding (DREB) TFs play a significant role in the regulation of gene expression in response to drought stress, activating the expression of genes involved in the biosynthesis of osmoprotectants and stress-responsive proteins. Similarly, abscisic acid (ABA)-responsive element-binding (AREB) TFs are involved in the regulation of gene expression in response to water stress, regulating the expression of genes involved in ABA signaling, stomatal closure, and stress-responsive genes [1]. One of the most important TFs in the regulation of basal resistance against invasive biotrophic and hemibiotrophic pathogens is the ENHANCED DISEASE SUSCEPTIBILITY 1 (EDS1) protein, which interacts with PHYTOALEXIN DEFICIENT 4 (PAD4) [2]. EDS1 is a plant-specific TF that coordinates the mobilization of basal immunity against various pathogenic strains recognized by specific NLRs, as well as effector-triggered immunity (ETI). These mechanisms are important in preventing pathogen infection and promoting plant survival [2]. Originally identified in a screen for mutants defective in resistance to isolates of the oomycete pathogen *Peronospora parasitica*, EDS1 and PAD4 also act as a regulatory hub for Toll-Interleukin-1 receptor (TIR)-type nucleotide binding-leucine rich repeat (NB-LRR) proteins to signal isolate-specific pathogen recognition [3]. Recent research highlights the fundamental role of EDS1 and PAD4 in transducing redox signals in response to biotic and abiotic stresses. EDS1 and PAD4 are important activators of salicylic acid (SA) signaling and antagonize the jasmonic acid (JA) and ethylene (ET) defense response pathways [4]. The EDS1-PAD4 heterodimers reinforce local and systemic defenses activated by various NLRs and contribute to basal immunity through transcriptional regulation. EDS1 family members demonstrate different modes of action and pathway connectivities that explain their central role in biotic stress resilience across seed plant lineages. Through these mechanisms, EDS1 and PAD4 play a fundamental role in biotic stress resilience across seed plant lineages [2].

However, TFs, particularly EDS1 and PAD4, play a critical role in regulating the plant's defense mechanisms against biotic and abiotic stresses. These proteins coordinate the mobilization of basal immunity, as well as effector-triggered immunity, and are involved in isolating pathogen recognition, transducing redox signals, and activating different defense response pathways. Their central role in biotic stress resilience across seed plant lineages demonstrates the significance of these proteins in plant survival and adaptation to their environments.

## MATERIALS AND METHODS

### Database Search and Retrieval of Sequences

The EDS domain amino acid sequence of EDS1 was obtained from Pfam-database, which was used in a BLAST-P (Protein- basic local alignment search tool) search on the cucumber genome database at Phytozome v13 (<https://phytozome-next.jgi.doe>) to identify CsEDS genes in cucumber [5].

The retrieved sequences were verified on NCBI CDD (Conserved Domain Database) (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) [6] with the default parameters. The proteins lacking real EDS domain (PF03106) (<https://pfam.xfam.org/family/PF03106>) in their sequences were excluded.

### Determination of Physio-Chemical Properties Of *Cseds1* Proteins

ProtParam tool (<http://web.expasy.org/protparam/>) was used to predict protein length (AA residues), molecular weight, and pI- value of *EDS1* proteins [7]. The information of gene IDs, chromosomal position, sequence of gene and protein, were retrieved from Phytozome. These *EDS1 domain-encoding* genes were renamed based on the order of their physical position.

### Conserved Motif and Domain Prediction Analysis

The analysis of motifs present in the proteins of CsEDS1 gene family was carried out using MEME suite under the default conditions, which included 20 motifs. The hit data obtained from NCBI CDD and Norwich phytogenic data from MEGA 11.0 were also utilized, and the resulting information was visualized using TBtools to display the conserved domains and motifs. Additionally, the exon and intron distribution of the CsEDS1 gene family was analyzed using the gene structure display server (GSDS) web tool, utilizing both the genomic and CDS sequences.

### Subcellular Localization and Gene Structure Analysis

Subcellular localizations of *CsEDS1* genes were predicted with the help of WoLF PSORT (<https://wolfpsort.hgc.jp/>) [8]. To determine the intron/exon arrangement of CsEDS1 genes, the genomic and coding sequences were obtained from Phytozome v13. Additionally, the gff3 file of the Cucumber genome was retrieved from the same source. Subsequently, these sequences were employed to visualize the gene structure using the Gene Structure Display Server (GSDS v2.0), which is a widely used tool for displaying and analyzing gene structures. The GSDS v2.0 was accessed through its web interface available at <http://gsds.cbi.pku.edu.cn/> [9].

### Comparative Phylogenetic Analysis

The AA sequences of CsEDS1 proteins were aligned using MUSCLE [10]. *CsEDS1* proteins sequence of *C. lanatus*, *C. sativus*, *C. maxima* and *A. thaliana* were used to generate the phylogenetic tree in MEGA 11 software with the neighbor joining (NJ) method and bootstrapping set at 1000 replications.

## Cis-Elements and Conserved Motifs Recognition

For the analysis of Cis-regulatory elements, 1000-bp upstream promoter sequence was retrieved from PlantCare database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>).

## Chromosomal Location, Gene Duplication and Synteny Analysis

Based on the information about the accessible cucumber genome on Phytozome, the TBtools programme displayed the chromosomal positions and duplications of *CsEDS1* genes. Divergence time of *CsEDS1* genes was estimated with the help of Ka/Ks values. Protein sequences alignment was made using MUSCLE and the number of Ka/Ks substitution rates were determined using Tbttools Software. [11]. Ka and Ks values were calculated for *CsEDS1* gene pair duplicates, using various modes of duplication. The Ka and Ks substitution rates were estimated using TBtools and the Simple Ka/Ks calculator option. The Ka/Ks ratio was used to determine the molecular evolutionary rates of each gene pair, while the time of divergence was estimated using the formula " $T = Ks/2\lambda$  ( $\lambda = 6.5 \times 10^{-9}$ )" representing neutral substitution. To exhibit the syntenic relationship of the *CsEDS1*, Micro Synteny view software in TBtools was used to construct a map in advance Circos feature.

## Transcriptome Analysis

Transcriptome analysis of *CsEDS1* genes were performed using three approaches i.e. Gene Expression Omnibus (GEO), Expression Atlas and Expressed sequence Tags (EST). Previously generated RNA-seq data was obtained from NCBI-GEO (<https://www.ncbi.nlm.nih.gov/geo/>) and Expression Atlas (<https://www.ebi.ac.uk/gxa/home>) of different conditions of temperature and photoperiodism, different stages of embryogenesis and in response to high temperature. For expression profiling, Reads Per Kilobases per Million mapped reads (RPKM) values from RNA-seq data were log<sub>2</sub> transformed [12].

## Mirna Analysis

The micro-RNA (miRNA) sequences related to *CsEDS1* genes in cucumber were identified from psRNATarget (<https://www.zhaolab.org/psRNATarget/>) using CDS sequences of all *CsEDS1* genes with default parameters [13]. Putative function of the identified miRNA was retrieved from previously performed In vitro and In vivo experiments.

# RESULTS

## Physiochemical Properties of *Eds1* Genes in Cucumber

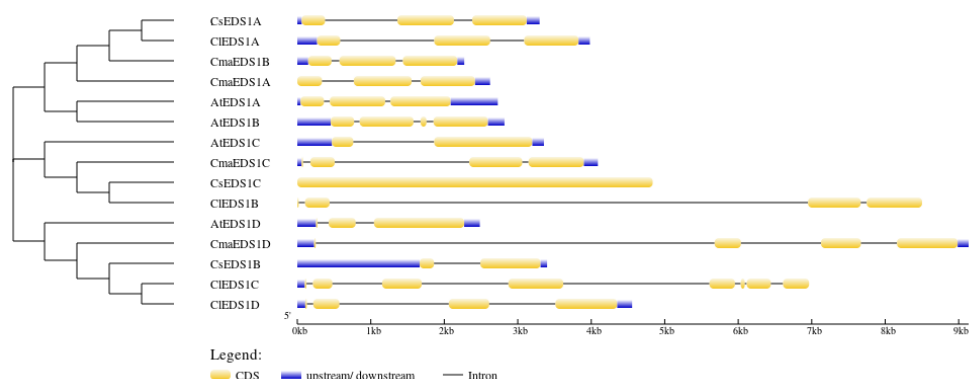
During BLAST-p analysis 3 *CsEDS1* genes in cucumber were identified and used for further analysis. The *CsEDS1* genes encoding protein's length vary from 337-608 AA and molecular weight vary from 69.78 to 39.45 kD with *CsEDS1C* being the smallest while *CsEDS1C* being the longest protein (Table 1).

**Table 1: Physiochemical properties of EDS1 genes in Cucumber**

Gene ID	Accession		Chromosome		Direction	Size (AA)		pI	Mw
Name	Phytozo me ID	Cucurbita	no.	Location (Base pairs)		mRNA	Peptid e		(KDa)
CsEDS1A	Cucsa.19 7280.1	CsaV3_1G 001870.1	scaffold 01357	723811..72 7107	F	1827	608	6.96	69.78
CsEDS1B	Cucsa.33 9400.2	CsaV3_3G 042050.1	3	2014028..2 017426	R	1014	337	5.4	39.45
CsEDS1C	Cucsa.35 3810.1	CsaV3_4G 030710.1	4	163144..16 7979	F	1791	596	7.09	68.04

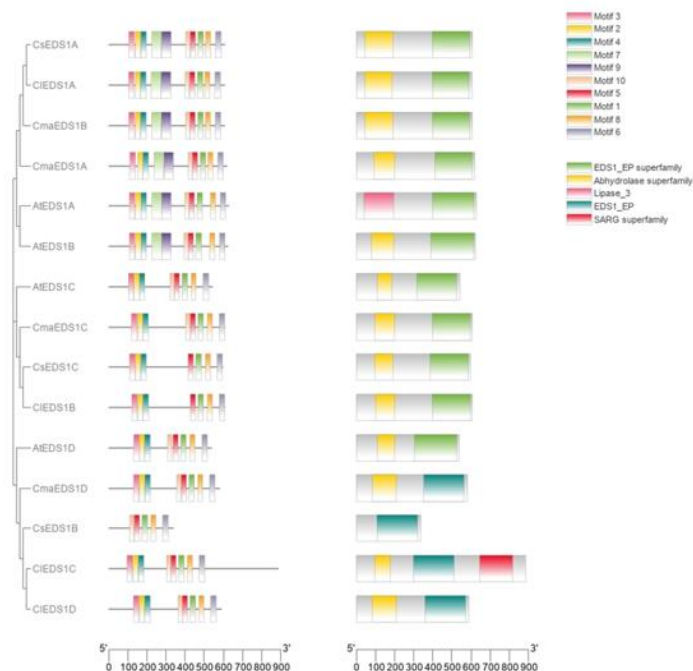
### Gene Structures & Recognition of Conserved Motifs and Domains

Exons and introns act as a backbone structures in determining the evolutionary relationship between genes [14]. Their numbers and distribution patterns are an evolutionary mark for a gene family. CsEDS1 had 2 introns, CsEDS1B was had 3 introns and CsEDS1C was intronless (Figure 1).



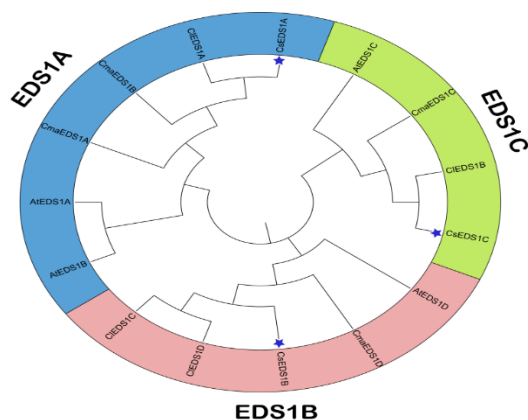
**Figure 1: Gene structure of CsEDS1 genes in Cucumber. Genomic and CDS sequences were used for drawing gene structure schematic diagrams with the Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/index.php>)**

All 15 motifs identified in the CsEDS1 *proteins* were studied using MEME program in Tbttools (Fig. 4). *EDS* domain was present in all the *CsEDS* proteins. It was observed that CsEDS1A has motif 3, 2, 4, 7, 9, 10, 5, 1, 8 and 6. CsEDS1C had motifs 3, 2, 4, 1, 8 and 6. CsEDS1B had motifs 10, 5, 1, 8 and 6 (Figure 2).



**Figure 2: Conserved Domain and Motif analysis of CsEDS1 genes in Cucumber**  
**Phylogenetic Analysis of *Eds1* Gene Family**

Evolutionary relationship between *C. sativus*, *A. thaliana*, *C. maxima* and *C. lanatu* were determined through Neighbor-Joining (NJ) phylogenetic tree. Results depicted that these total 15 eDS1 genes were divided in 3 clades, I, II and III. Clade I included AtEDS1C, CmaEDS1C, ClEDS1B and CsEDS1C. Clade II included AtEDS1D, CmaEDS1D, CsEDS1B, ClEDS1D and ClEDS1C. Clade III included AtEDS1B, AtEDS1A, CmaEDS1A, CmaEDS1B, ClEDS1A and CsEDS1A. Proteins in common clade mostly have similar structure and function (Figure 3). So, these *EDS1* proteins may have similar structures and functions.



**Figure 3: Phylogenetic Tree analysis of EDS1 genes in *C. sativus*, *A. thaliana*, *C. maxima* and *C. lanatus***

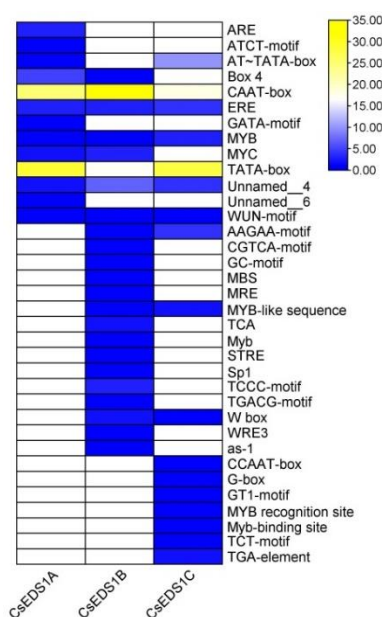
## Cis-Regulatory Elements and Subcellular Localization Analysis

Cis-regulatory elements affect the transcriptomic expression at the binding site of TrF's on the promoter region. In silico analysis can be performed to evaluate the putative function of various cis-elements in genes [15,16].

**S Table 1: Subcellular localization of CsEDS1 genes in Cucumber**

Gene Name	Subcellular Localization
CsEDS1A	Cytoplasm, Nucleus, Chloroplast, Mitochondria, Plastid
CsEDS1C	Chloroplast, Cytoplasm, Nucleus, Plastid
CsEDS1B	Cytoplasm, Chloroplast, Cysk

CAAT box and TATA box were found in all three CsEDS1 sequences. Except for that ERE, MYB, MYC, WUN motif were found in almost all the sequences of CsEDS1. These cis regulatory elements are involved in wound response, drought tolerance and hormone response (Figure 4). All of the EDS sequences were found to be localized in nucleus, cytoplasm and chloroplast (S Table 1).



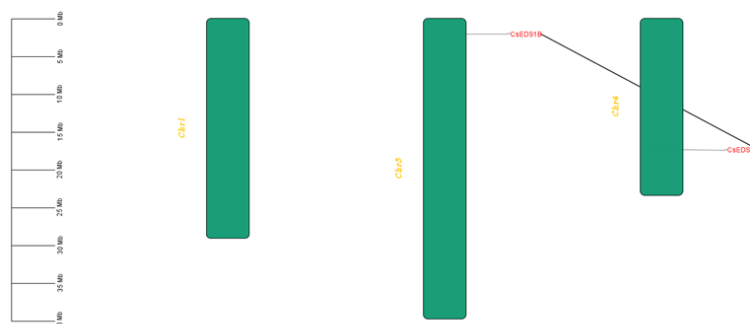
**Figure 4: Cis-regulatory elements of CsEDS1 genes in Cucumber**

## Chromosomal Mapping, Gene Duplication and Synteny Analysis

CsEDS1B on chromosome 3 is segmentally duplicated on CsEDS1C on Chromosome 4 (Table 2). Generally, the  $K_a/K_s < 1$  infers evolution as a purifying selection.  $K_a/K_s = 1$  implies an evolution to be a neutral selection, whereas  $K_a/K_s > 1$  is indicative of a positive selection. The ratio for  $K_a/K_s$  was found for all the cucumber pairs. It was seen that  $K_a/K_s$  ratio was less than 1 for all the pairs, which mean that they are the result of purification selection by environment (Figure 5).

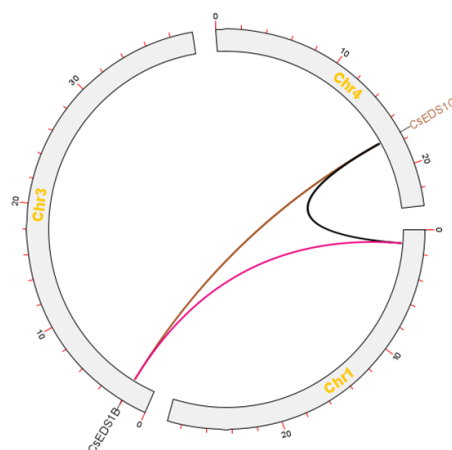
**Table 2: Gene Duplication Analysis of CsEDS1 genes**

Seq 1_Seq 2	Ka	Ks	Ka_Ks	MYA
CsEDS1A_CsEDS1B	0.741966	NaN	NaN	NaN
CsEDS1A_CsEDS1C	0.744765	NaN	NaN	NaN
CsEDS1B_CsEDS1C	0.975366	2.608229	0.373957	200.6330237



**Figure 5: Chromosomal Mapping of CsEDS1 genes**

Synteny Analysis for CsEDS1 proteins showed 100% segmental duplication. CsEDS1C on chromosome 4 is segmentally duplicated on chromosome 1. CsEDS1B on chromosome 3 is segmentally duplicated on chromosome 1 and 4 (Figure 6).



**Figure 6: Synteny Analysis of CsEDS1 genes to find paralogues genes**

## Transcriptome Analysis

### Embryogenesis

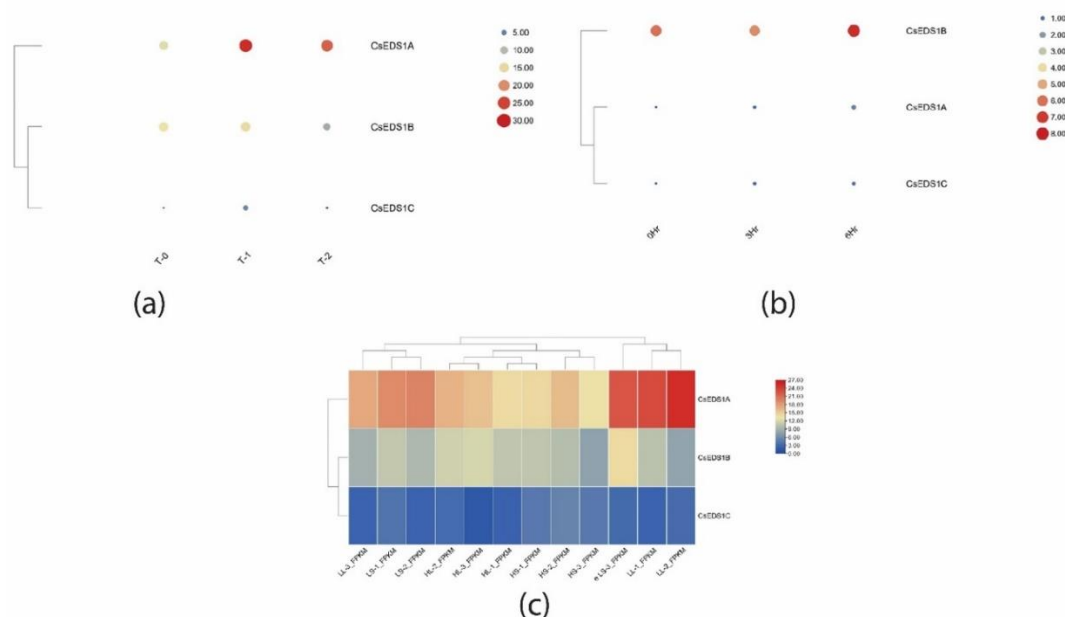
The expression of CsEDS1A was seen to be overexpressed after T-1 and T2 but the overexpression isn't significant. The expression of CsEDS1B and CsEDS1C was seen to be downregulated significantly after T1 and T2 (Figure 7(a)).

## High Temperature Stress

The expression of CsEDS1A was seen to be increased after the high temperature treatment whereas the expression of CsEDS1B and CsEDS1C was seen to be downregulated after high temperature treatment. The downregulation of CsEDS1C was seen to be significant (Figure 7(b)).

## Different Conditions Of Photoperiodism And Temperature

The expression of CsEDS1A was seen to be upregulated in response to long day and short dark period whereas the expression of CsEDS1B and CsEDS1C was seen to be downregulated. The overexpression of CsEDS1A and downregulation of CsEDS1C was seen to be significant (Figure 7(c)).



**Figure 7: (a) Transcriptome Analysis with respect to embryogenesis (b) Transcriptome Analysis with respect to high temperature stress (c) Transcriptome Analysis with respect to different conditions of Photoperiodism and Heat**

**S Table 2: Putative miRNA analysis of CsEDS1 genes in Cucumber**

miRNA Acc.	Target Acc.	Expect	UPE	Alignment	Inhibition	Multiplicity
csa-miR7499	CsEDS1C	3.0	N/A	miRNA 23 UUGGGUUUAUUGGGCUUUUAUAU 1 ..... Target 785 UGUAUAAAUAGCCUAAAAUAUA 807	Cleavage	1
csa-novel-mir23	CsEDS1C	3.5	N/A	miRNA 21 ACGAAACAAUAGAUGAAAAAU 1 .....	Cleavage	3

				Target AUUUUUGCUAUUUACCUUUUA 1674	1654		
csa-novel-mir23	CsEDS1C	3.5	N/A	miRNA ACGAAACAAUAGAUGAAAAU 1 : : : : : : : : : : Target CUUUCUUAUUUUUACUUUUUG 544	21 524	Cleavage	3
csa-novel-mir230	CsEDS1C	3.5	N/A	miRNA UUUUGCUAGUAAAAUUUAAAUC 1 : : : : : : : : : : Target AUAACAAACAUUUGAAUUUAU 1936	21 1916	Cleavage	2
csa-novel-mir132	CsEDS1C	4.0	N/A	miRNA GAAGAAGAAAAGAGAGGCAGAAA 1 : : : : : : : : : : Target UUUGUUUUUUUUUUUUUCUUUUUU 633	23 611	Cleavage	3
csa-novel-mir145	CsEDS1C	4.0	N/A	miRNA UCAGAUGACAAAGUAAGAACU 1 : : : : : : : : : : Target UUUUUACUCUUUCAUUUUUUA 460 480	21 460 480	Cleavage	1
csa-novel-mir23	CsEDS1C	4.0	N/A	miRNA ACGAAACAAUAGAUGAAAAU 1 : : : : : : : : : : Target UGUUUCGUUAUAUACUUUUUU 1061	21 1041	Translation	3
csa-novel-mir257	CsEDS1C	4.0	N/A	miRNA AUAUGAGGAAGGUGAAGAGAGUA 1 : : : : : : : : : : Target ACUAAUCAUUUCAUUUUUUUCAU 460	23 438	Cleavage	1
csa-novel-mir47	CsEDS1A	4.0	N/A	miRNA AUUUUUGAUUUUACGUCGAGUU 1 : : : : : : : : : : Target UGAAUCUAAAAUGCAGCAAAA 960 980	21 960 980	Cleavage	1
csa-miR167	CsEDS1C	4.5	N/A	miRNA AUCUAGUACGACCGUCGAAGU 1 : : : : : : : : : : Target ACGAUAGUGUUUGUAGUUUCA 686	21 666	Translation	1
csa-miR2628	CsEDS1C	4.5	N/A	miRNA AAUAAGUAGAAGAGAAAGUC 1 : : : : : : : : : : Target GUGUUCUUUUUGUCUUUUUGU 596 615	20 596 615	Cleavage	2
csa-miR5747	CsEDS1C	4.5	N/A	miRNA AUUACAUAUACUCAUAAGAUAU 1	23	Cleavage	1

				Target UUCUUUUUUUAUGAGCAUUCUAA 1082		
csa-miR8125	CsEDS1C	4.5	N/A	miRNA AUGGAAAAGUGUAAGAAGAGGAC 1 Target UUUUUUUUUUUCUUUUUUUUUUUG 639	Cleavage	1
csa-miR8693	CsEDS1C	4.5	N/A	miRNA UACGAGUAGAUAGAGAAAGUAGA 1 Target AUUAUCAUCUCUUUUUUUUUAU 1343	Cleavage	3
csa-novel-mir132	CsEDS1C	4.5	N/A	miRNA GAAGAAGAAAAGAGAGGCAGAAA 1 Target UUAUUUCUUCUCUUUCCUUUUUC 2537	Cleavage	3
csa-novel-mir21	CsEDS1C	4.5	N/A	miRNA ACGUGAAGAGAUUUGGGAUUU 1 Target CCUAAUUCUUUAAAUCUUUAA 361	Cleavage	2
csa-novel-mir223	CsEDS1C	4.5	N/A	miRNA UGGCUGUUGUAAUAUGUUUUCAA 1 Target AACCACGACAUUAUGCCAAGGCU 3534	Cleavage	1
csa-novel-mir288	CsEDS1C	4.5	N/A	miRNA UUGGAAGUAAGUUGAAUGGAGAU 1 Target AGUUUUCUAAGACUUAUCUUUG 2846	Cleavage	2
csa-novel-mir29	CsEDS1C	4.5	N/A	miRNA AUUUGAUGAUAAAGUAAAGGUU 1 Target GUAAUUGUAAUUUUUUUUUAA 2090	Cleavage	1
csa-novel-mir309	CsEDS1C	4.5	N/A	miRNA UAAUAUUGAAGAUGAACAAAGAU 1 Target CUUCUAAUUUUUUUUUUUCUUUCUA 530	Translation	1

## miRNA Analysis

Studying Putative Mirna Targets in genome-wide analysis is crucial for understanding the regulatory role of miRNAs in gene expression, and predicting and validating miRNA targets can provide valuable insights into disease mechanisms and guide the development of therapeutic interventions. In our research, 22 specific miRNAs were identified, ranging from 20 to 23 amino acids in length, that exclusively target CsEDS1C, with some inhibiting translation and others inhibiting cleavage of the targeted gene. Notably, multiple miRNAs were found to target a single gene, indicating potential coordinated regulation, and the inhibition of CsEDS1C expression by these miRNAs may suggest a broader impact on mRNA degradation (S Table 2).

## DISCUSSION

Plant intracellular receptors known as nucleotide binding–leucine-rich repeat (NB-LRR) receptors are capable of intercepting pathogen effectors, but the mechanism by which these receptors activate downstream defenses is still unclear. EDS1, a nucleocytoplasmic basal resistance regulator, is crucial for Toll–interleukin-1 receptor (TIR)–NB-LRR receptor-mediated immunity. While nuclear processes can restrict bacterial growth, coordinated nucleo-cytoplasmic activity is necessary for both programmed cell death and transcriptional resistance reinforcement. As a result, EDS1 functions as both an effector target and an activated TIR-NB-LRR signal transducer, providing defenses across cell compartments [17].

To investigate the role of EDS1 protein in plant defense in *Cucumis sativus*, the protein sequences were retrieved and various structural and functional analyses were performed to trace their role in plant defense mechanism. The results of all the structural and functional analyses were also compared with *Arabidopsis thaliana*, *Citrullus lanatus* and *Cucurbita maxima* for further validation a phylogenetic tree was constructed which divided all these sequences in three clades named as Clade I contains AtEDS1C, CmaEDS1C, CIEDS1B, and CsEDS1C, while Clade II includes AtEDS1D, CmaEDS1D, CsEDS1B, CIEDS1D, and CIEDS1D. Clade III consists of AtEDS1A, AtEDS1B, CmaEDS1A, CmaEDS1B, CIEDS1A, and CsEDS1A. The CsEDS1 proteins, which are in the same clade as AtEDS1, have identical functions. To understand the basic structure of CsEDS1 proteins, the EDS1 domain was found in all the CsEDS1 sequences. Gene structure intron exon display showed the slight variation in the number of introns in various CsEDS1 sequences. Similarly, the occurrence of motifs showed slight variation which indicates the conservation of most of the structure with slight expansion. Another evidence of genetic expansion is the 100% segmental duplication found in chromosomal mapping and synteny analysis. The various cis-regulatory elements, such as ERE, WUN motif, MYB, and MYC, present in almost all CsEDS1 proteins have been shown to play a crucial role in several processes including wound healing, hormone signaling, defense pathways, and drought resistance. The presence of these transcription factors (TFs) in the promoter region of CsEDS1 proteins indicates their involvement in plant resistance. To delve deeper into their functional significance, the expression of various CsEDS1 proteins was analyzed in silico using NCBI GEO. The analysis revealed that CsEDS1C exhibited

significant downregulation during embryogenesis, thus confirming its role in this process. Additionally, CsEDS1C was also significantly downregulated in response to high temperature stress, suggesting its involvement in heat stress response. On the other hand, the overexpression of CsEDS1A and downregulation of CsEDS1C indicated their significant roles in different conditions of photoperiodism and temperature. Thus, it can be inferred that CsEDS1 plays a significant role in embryogenesis, heat stress response, and also responds to day length and temperature.

## CONCLUSION

The study investigated the role of EDS1 protein in plant defense mechanism in *Cucumis sativus*. Various structural and functional analyses were performed, and the results were compared with other plant species for validation. A phylogenetic tree was constructed, and CsEDS1 proteins were divided into three clades. CsEDS1 proteins were found to have the EDS1 domain, and slight variations were observed in gene structure and motifs. The study also revealed that CsEDS1 proteins are involved in plant resistance and respond to various environmental conditions. Overall, CsEDS1 is significantly involved in embryogenesis, heat stress, and photoperiodism, and it plays a crucial role in plant defense. This study also provides valuable insights into their molecular characteristics, evolutionary relationships, and regulatory functions. These findings can aid in the development of crop improvement strategies to enhance plant resistance to biotic and abiotic stresses, ultimately benefiting agriculture and food security.

## Conflict of Interest

The author declares that there is no conflict of interest.

## Author's Contribution

MH and AS conceived and conceptualized the study and wrote original draft. MS and M. ZH technically reviewed and finalized the draft.

## References

1. Schwechheimer, C.; Zourelidou, M.; Bevan, M. Plant transcription factor studies. *Annual review of plant biology* **1998**, *49*, 127-150.
2. Wiermer, M.; Feys, B.J.; Parker, J.E. Plant immunity: the EDS1 regulatory node. *Current opinion in plant biology* **2005**, *8*, 383-389.
3. Parker, J.E.; Holub, E.B.; Frost, L.N.; Falk, A.; Gunn, N.D.; Daniels, M.J. Characterization of eds1, a mutation in *Arabidopsis* suppressing resistance to *Peronospora parasitica* specified by several different RPP genes. *The Plant Cell* **1996**, *8*, 2033-2046.
4. Lapin, D.; Bhandari, D.D.; Parker, J.E. Origins and immunity networking functions of EDS1 family proteins. *Annual review of phytopathology* **2020**, *58*, 253-276.
5. Rinerson, C.I.; Rabara, R.C.; Tripathi, P.; Shen, Q.J.; Rushton, P.J. The evolution of WRKY transcription factors. *BMC plant biology* **2015**, *15*, 1-18.
6. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* **2020**, *48*, D265-D268, doi:10.1093/nar/gkz991.

7. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook*; Springer: 2005; pp. 571-607.
8. Horton, P.; Park, K.-J.; Obayashi, T.; Nakai, K. Protein subcellular localization prediction with WoLF PSORT. In Proceedings of the Proceedings of the 4th Asia-Pacific bioinformatics conference, 2006; pp. 39-48.
9. Hu, B.; Jin, J.; Guo, A.Y.; Zhang, H.; Luo, J.; Gao, G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* **2015**, *31*, 1296-1297, doi:10.1093/bioinformatics/btu817.
10. Edgar, R.C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **2004**, *5*, 113, doi:10.1186/1471-2105-5-113.
11. Li, G.; Hou, M.; Liu, Y.; Pei, Y.; Ye, M.; Zhou, Y.; Huang, C.; Zhao, Y.; Ma, H. Genome-wide identification, characterization and expression analysis of the non-specific lipid transfer proteins in potato. *BMC Genomics* **2019**, *20*, 375, doi:10.1186/s12864-019-5698-x.
12. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant* **2020**, *13*, 1194-1202, doi:https://doi.org/10.1016/j.molp.2020.06.009.
13. Samad, A.F.A. MicroRNA and Transcription Factor: Key Players in Plant Regulatory Network. *Frontiers in Plant Science* **2017**.
14. Koralewski, T.E.; Krutovsky, K.V. Evolution of exon-intron structure and alternative splicing. *PloS one* **2011**, *6*, e18055.
15. Bulow, L.; Hehl, R. Bioinformatic Identification of Conserved Cis-Sequences in Coregulated Genes. *Methods Mol Biol* **2016**, *1482*, 233-245, doi:10.1007/978-1-4939-6396-6\_15.
16. Jones, D.M.; Vandepoele, K. Identification and evolution of gene regulatory networks: insights from comparative studies in plants. *Curr Opin Plant Biol* **2020**, *54*, 42-48, doi:10.1016/j.pbi.2019.12.008.
17. Heidrich, K.; Wirthmueller, L.; Tasset, C.; Pouzet, C.; Deslandes, L.; Parker, J.E. Arabidopsis EDS1 connects pathogen effector recognition to cell compartment-specific immune responses. *Science* **2011**, *334*, 1401-1404.