Identification and comprehensive analyses of the *CBL* gene family in sweet orange (*Citrus sinensis* L.)

Original paper

Seyyed Hamidreza Hashemipetroudi [1,*]**, Hamidreza Ghorbani** [2]**, Firozeh Sohrevardi**[1]**, Mozhdeh arab**[1,3]

[1.] Genetic and Agricultural Biotechnology Institute of Tabarestan (GABIT), Sari Agricultural Sciences and Natural Resources University (SANRU), Sari, Iran
[2.] Crop and Horticultural Science Research Department, Mazandaran Agricultural and Natural Resources Research and Education Center, AREEO, Sari, Iran
[3.] National Institute of Genetic Engineering and Biotechnology (NIGEB), Tehran, Iran

Abstract: Calcineurin B-like (CBL) proteins, as calcium sensors, serve roles in plant responses to varied abiotic stressors and in growth and development through interaction with CBL-interacting protein kinases (CIPKs). However, information on the roles and development of CBLs in sweet orange plants is limited. We surveyed the whole *Citrus sinensis* genome and found eight *CBL* genes. Domains features, position and distribution, and conserved motif revealed that the EF-hands domain was conserved across the eight CsCBLs. CsCBL proteins are classed as acidic CBL, and five myristoylation sites and six palmitoylation sites were predicted. Eight *CsCBLs* were distributed across chromosomes Chr01, Chr02, Chr04, and Chr05 and contig chrNW-006257104.1. In chromosome 05, tandem duplications likely gave rise to two *CsCBL4* and *CsCBL5* genes. The phylogeny tree of 37 CBL proteins from different plant species including *Arabidopsis thaliana*, *Oryza sativa*, *Sesamum indicum*, and *C. sinensis* showed that these CBLs are closely related. A meta-analysis of the *CsCBL* gene family's expression in different tissues/stresses revealed that *CsCBL* genes expressed differently in tissues, which could be evidence for *CsCBL* tissue/stress-specific expression. The results of this study highlight the functional properties of the *Cs*CBL gene family and provide crucial data for future research on their functional activities.

Keywords: calcineurin B-like, Calcium sensor, CBL, *Citrus sinensis*, Sweet orange, signaling, stress.

Introduction

There are many biological systems that depend on the calcium ion to regulate basic growth and development processes and respond to diverse environmental challenges (Luan, 2009; Ma et al., 2020). Depending on the kind of cell, the concentration of Ca^{2+} might range from nM to mM (Tuteja and Mahajan, 2007; Mohanta and Sinha, 2016). Several Ca^{2+} sensors, including calcium-dependent protein kinases (CDPKs), calcium modulin, and calcineurin B-like proteins (CBLs), detect the Ca^{2+} signals, which are then sent to downstream destinations, triggering a variety of physiological reactions in the cell (Kanchiswamy et al., 2013; Sarwat et al., 2013). Each CBL contains at least three EF domains and Ca^{2+}-binding sites (Mao et al., 2016). By interacting with the CBL-interacting protein kinases (CIPKs), CBLs transfer Ca^{2+} signals (Ma et al., 2020; Tang et al., 2020). CBLs contain EF-hands and myristoylation sites, which are both conserved domains (Kolukisaoglu et al., 2004). The EF-hands are calcium sensors that may be identified by the presence of a conserved Asp (D) or Glu (E) residue (Kanchiswamy et al., 2013). Additionally, each of the themes has a helix-loop-helix structure with a total of 36 amino acid residues. Calmodulines and CBLs, in contrast to CDPKs, are small proteins that do not have an effector kinase domain.

Plant responses to environmental stressors such as drought, salinity, and lack or excess of nutrients can be controlled by CBL-CIPK (Sanyal et al., 2015; Thoday-Kennedy et al., 2015). In addition to this, it regulates the growth and development of plants, as well as the absorption and/or transport of nitrate, ammonium, and iron; the maintenance of H^+ homeostasis; and the transduction of signals involving reactive oxygen species (Li et al., 2016; Mao et al., 2016). Ten genes related to CBL proteins (CBL1-10) have been found in *Arabidopsis* with similar structural domains but minor differences in coding region length (Kolukisaoglu et al., 2004). CBL1 and CBL9 were shown to control the transmission of K^+, nitrate, and ammonium, as well as the development of stomatal opening and closing and ROS signaling (Li et al., 2006; Xu et al., 2006; Lu et al., 2017; Yin et al., 2017). Pollen germination and tube expansion are controlled by CBL2 and CBL3, which sequester Mg^{2+} (Steinhorst et al., 2015).

Additionally, CBL3 is responsible for the distribution and translocation of K^+ (Liu et al., 2013). CBL10 is involved in fruit growth and quality, as well as salt tolerance, K^+ absorption, and GTPase activity modification (Kim et al., 2007; Chow et al., 2016).

CBL gene family members have been discovered at the genome-wide level in rice, maize, wheat, and other plants in recent years. CBLs are also critical regulators of plant responses to diverse abiotic stresses, as well as growth and development (Yu et al., 2014; Thoday-Kennedy et al., 2015). Cotton (*Gossypium hirsutum*) fiber elongation appears to be modulated by GhCBL2 and GhCBL3 (Gao et al., 2008). Ten *OsCBL* genes in rice have been reported to respond to diverse stress conditions [sodium chloride (NaCl), polyethylene glycol (PEG), and cold] and are expressed in several organs in the adult stage. In addition, *OsCBL8* overexpressing transgenic rice seedlings exhibited greater salt tolerance than non-transgenic seedlings (Gu et al., 2008). Cellular adaption responses to sodium carbonate stress in *S. bicolor* are assumed to be controlled by *CBL* genes, which are known to govern plant growth and development patterns (Zhang et al., 2011). Under abiotic stress, the transcripts of seven *CBL* members (PeCBL1 to PeCBL10) were found to play key roles in the response of *Populus euphratica* to certain external stimuli (Zhang et al., 2008). In their comprehensive study on the calcium sensor families in the halophyte plant *Aeluropus littoralis*, Arab et al. (2023) revealed the responses of the CBL-CIPK network to salt stress. In the current study, the CBL gene family in sweet orange was identified and characterized based on the newly released genome of *Citrus sinensis* (Wu et al., 2018). Genome-wide characterization of *CiCBL* was analyzed using the available bioinformatics tools, which included gene structure, protein domain, phylogenetic and evolutionary approaches, and gene expression profiling, to better understand the evolutionary history of *CBL* genes.

Materials and Methods

CsCBL searching and characteristics

The complete genome assembly of citrus (*C. sinensis* 'Valencia') was downloaded from the Citrus

Genome Database (www.citrusgenomedb.org; *C. sinensis* genome v2.0). Ten CBL protein sequences from *A. thaliana* (AtCBLs) were obtained from the *Arabidopsis* Information Resource (TAIR) database (www.arabidopsis.org/). Candidate citrus CBL protein sequences (CsCBLs) were selected by blasting the AtCBLs as the query sequences in proteome sequence of *C. sinensis* via local BLASTP. The default statistical parameters used in the BLASTP analysis were as follows: BLASTP-protein query of the protein database; expected threshold (E): −1, comparison matrix: BLOSUM62; no. of alignments to show: 100. Redundant proteins were manually deleted based on their E-values and identity. The identified CBLs were then checked using the Pfam database (www.pfam.xfam.org), and conserved EF-hand domains were validated for all putative CBL proteins using the NCBI CD-Search program. InterProScan (www.ebi.ac.uk/inter pro/) was used to perform the domain analysis against the protein database. ExPASy (wwww.expasy.org/compute_pi/) was used to calculate the candidate protein molecular weight (MW) and isoelectric point (pI). The MEME website (http://meme-suite.org/tools/meme) was used to find conserved motifs with the following optimized parameters: zero or one occurrence per sequence, a maximum of 10 motifs, and an optimum motif width of 6 to 50 residues. For all other parameters, default values were applied.

Phylogenetic and gene structure analysis

The 37 CBL protein sequences from *A. thaliana*, *Oryza sativa*, *Sesamum indicum* and *C. sinensis* were aligned via muscle in MEGA version 7.0, with default parameters (Tamura et al., 2013). In addition to that, a tree constructed using the neighbor-joining (NJ) algorithm was produced through bootstrapping (1000 replicates). The structure of *CsCBLs* was discovered using GFF information by matching the coding sequences to the appropriate genomic sequences. In addition, MEME program was used to create an illustration of the CsCBL protein motifs, conserved domain, gene architectures, and a phylogenetic tree. The chromosome locations of the candidate *CsCBL* genes were investigated by MCScanX (Wang et al., 2012).

Expression profiling of CsCBL genes based on RNA-seq data

To examine the function and expression of the *CsCBL* gene family, 427 RNA-seq samples from 18 bioprojects were gathered from publicly accessible RNA-seq data associated with *C. sinensis*. Transcriptome datasets were selected from various treatments or genetic backgrounds, their accession number were listed as follow: PRJNA741128, PRJNA715742, PRJNA704425, PRJNA703546, PRJNA674975, PRJNA612768, PRJNA566421, PRJNA517400, PRJNA488876, PRJNA471083, PRJNA386941, PRJNA350382, PRJNA340305, PRJNA339838, PRJNA203307. TPM (transcripts per million) had been utilized to measure transcript expression levels.

Results

Eight distinct *CsCBL* were found in *C. sinensis* (Table 1 and Supplementary Table S1). All *CsCBL* have 7 introns and 8 exons, except *SsCBL7*, which has 8 introns and 9 exons. They all code for between 213 (CsCBL3) and 259 amino acid residues (CsCBL8). For the most majority of CBLs, the theoretical isoelectric point (pI) is low and less than five, ranging from 4.67 (CsCBL3) to 5.07 (CsCBL3 and CsCBL4). Because their pI value is less than 6, plant CBL proteins are classed as acidic CBL proteins. The range of molecular weight was between 24.421 (CsCBL3) to 29.671 KDa (CsCBL8). There were five myristoylation sites and six palmitoylation sites predicted in the CsCBLs. Besides, myristoylation sites were not found on CsCBL2, CsCBL7, and CsCBL8, and palmitoylation sites were absent in CsCBL7 and CsCBL8 (Table 1). In the N-terminal region of three CsCBLs (CsCBL3, CsCBL5, and CsCBL6), conserved myristoylation and palmitoylation sites (M-G-C-) can be found. These sites play important roles in protein aggregation, stability, and trafficking.

The proteins' stability indexes are shown in Table 1. According to the evidence presented, CsCBL3 (36.71), which provides the most unstable CsCBL, and CsCBL2 (46.16), which provides the most stable CsCBLs, respectively. A peptide's gravy value is computed by multiplying the hydropathy values of each residue by the length of the sequence. The gravy of CsCBL8 (0.044) is more hydrophilic than

the other CsCBLs, according to Table 1. CsCBL1 also has a lower hydrophilicity than the other CsCBL.

Distribution of the CsCBL gene family members in the C. sinensis genome

In this study, the chromosomal distributions of the *CBL* genes were investigated in *C. sinensis* genome. In total, there were eight *CsCBLs* on five chromosomes (Chr01, Chr02, Chr04, Chr05 and chrNW-006257104.1). Cschr05 included three *CsCBLs*, while Cschr04 contained two *CsCBLs*. Cschr01, Cschr02, and CschrNW-006257104.1 each had one *CsCBL*. The positions of *CsCBL* gene family on the chromosomes of *C. sinensis* are shown in Figure 1. There were two *CBL* genes (*CsCBL4* and *CsCBL5*) identified in Cschr05 that were found to be linked together in a chromosome. Tightly linked genes are likely to have been created through tandem duplications, according to these findings. Both the duplication of genes and the divergence of their functions play an important role in the expansion of gene families and the development of new activities.

Gene structure, domains and motifs arrangement of the CsCBLs

CsCBLs were identified based on their EF-hand sites and number of EF-hand domain. All of the CBL members have been found to have EF-hand motifs, which are responsible for binding to Ca^{2+} ions and transferring calcium signals. All sequences contained four EF-hand motifs.

All CBL proteins contained EF-hand 7, and other EF-hands were located on various CBL proteins (Figure 2a). MEME motif search was used to uncover conserved motifs in 8 CsCBL proteins (Figure 2b) since CBL domains are critical for the action of CBLs. The first two and third motifs were the longest of the ten found motifs (Figure 3 and Table 2). As shown in Figure 2b, motifs 1-5 were found in all of the CsCBLs. In addition, when compared to the other CsCBLs, CBL5 and CBL8 were missing motif 9, whereas CBL1, CBL2, and CBL8 were missing motif 7. Motifs 10 and 8 were only present in CBL7 and CBL8, while motif 6 was found in CBL1 and CBL2. As shown in Figure 2b, motifs 1-5 were found in all of the CsCBLs. In addition, when compared to the other CsCBLs, CBL5 and CBL8 were missing motif 9, whereas CBL1, CBL2, and CBL8 were missing motif 7.

Motifs 10 and 8 were only present in CBL7 and CBL8, while motif 6 was found in CBL1 and CBL2.The pattern with the maximum number of motifs was found in CBL7, which did not have motif 6 (contains nine motifs), while the pattern with the lowest number of motifs was found in CBL5, which did not contain motif 6, 8, 9, or 10 (six motifs) (Figure 2b). *CsCBL* gene family had a different number of introns (between 7 and 9), and the UTR-CDS structure and the length of the coding regions were different between *CsCBLs* (Figure 2c).

**Table 1.** Physicochemical characteristics of *C. sinensis* CBL gene family.

Gene ID	Molecular weight (KDa)	Isoelectric point	Instability index	Aliphatic index	Gravy	N-Myristoylation	Palmitoylation
CsCBL1	26.057	4.70	39.86	90.97	-0.265	*MVQCLDGLKHFCVV	****MVQCLDGLKHF LDGLKHFCVVVVNCC FCVVVVNCCDADLYK CVVVVNCCDADLYKQ
CsCBL2	25.668	4.81	46.16	90.94	-.239	-	LFASLLQCCDTNPSR
CsCBL3	24.421	4.67	36.71	91.97	-0.114	******MGCFQSKVA	*****MGCFQSKVAK
CsCBL4	24.495	5.07	38.70	93.41	-0.136	***MNAFGRCFCMKK	*MNAFGRCFCMKKSK NAFGRCFCMKKSKQI
CsCBL5	24.839	4.86	43.57	104.68	-0.025	******MGCVCMKQR	*****MGCVCMKQRL ***MGCVCMKQRLKS
CsCBL6	25.152	4.77	41.73	90.82	-0.187	******MGCVLTKRT	****MGCVLTKRTK
CsCBL7	29.000	4.58	38.08	87.18	-0.101	-	-
CsCBL8	29.671	4.73	45.47	100.97	0.044	-	-
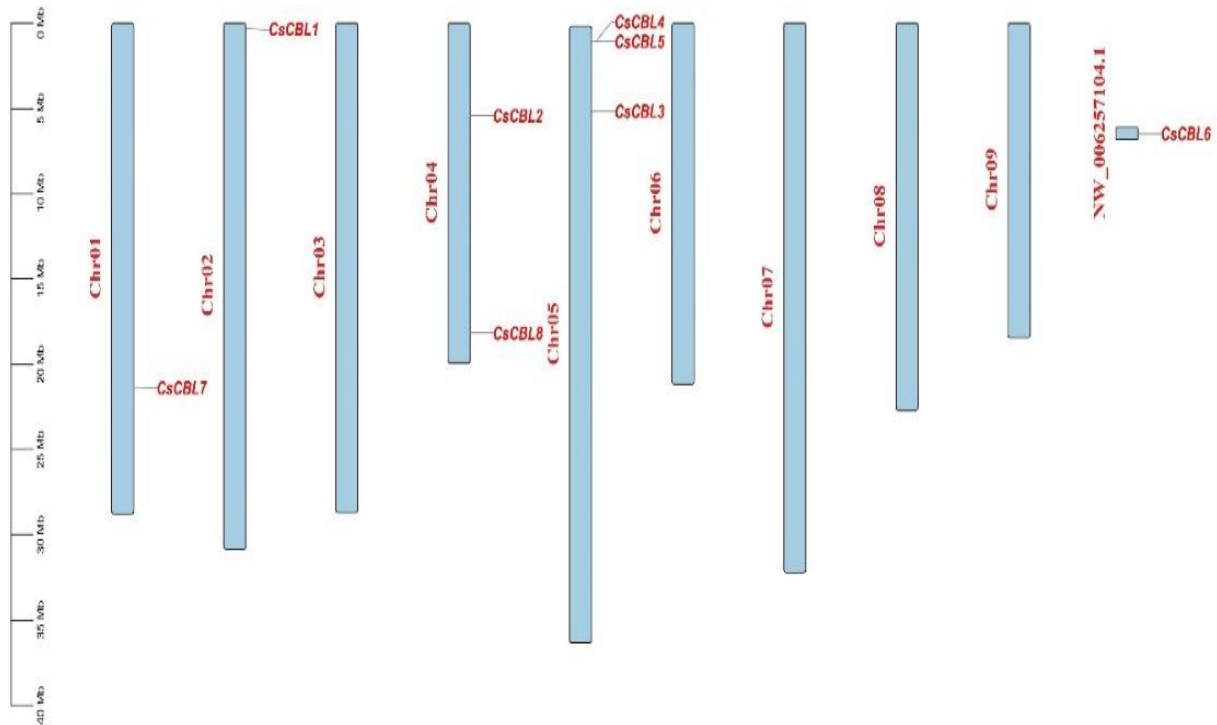
Figure 1. The chromosomal distribution of the *CBL* gene family in *Citrus sinensis*. The left side of each chromosome displays the chromosome numbers and their respective approximate sizes.

Table 2. Sequence of identified motif in eight CBL family members of *C. sinensis*.

Motif ID	Sequence	WIDTH
MEME-1	IIDKTFEDADTKGDGKIDKEEWKEFVLRNPSLLKNMTLPYLKDITTAFPS	50
MEME-2	RNGVIEFEEFVRALSIFHPNAPIEDKIDFAFRLYDLRQTGFIEREEVKQM	50
MEME-3	EIEALYELFKKJSSSIIDDGLIHKEEFQLALFKNSKKENLFADRVFDLFD	50
MEME-4	KQRPGYEDPVILAAETPFSVS	21
MEME-5	VVALLKESEMKLSDD	15
MEME-6	MLQCJEGFKHFCVVLLNCCD	20
MEME-7	MGCFCT	6
MEME-8	YWGSSSLQFGEKJCAVCIP	19
MEME-9	FVFHSEVD	8
MEME-10	MDSAAN	6

1
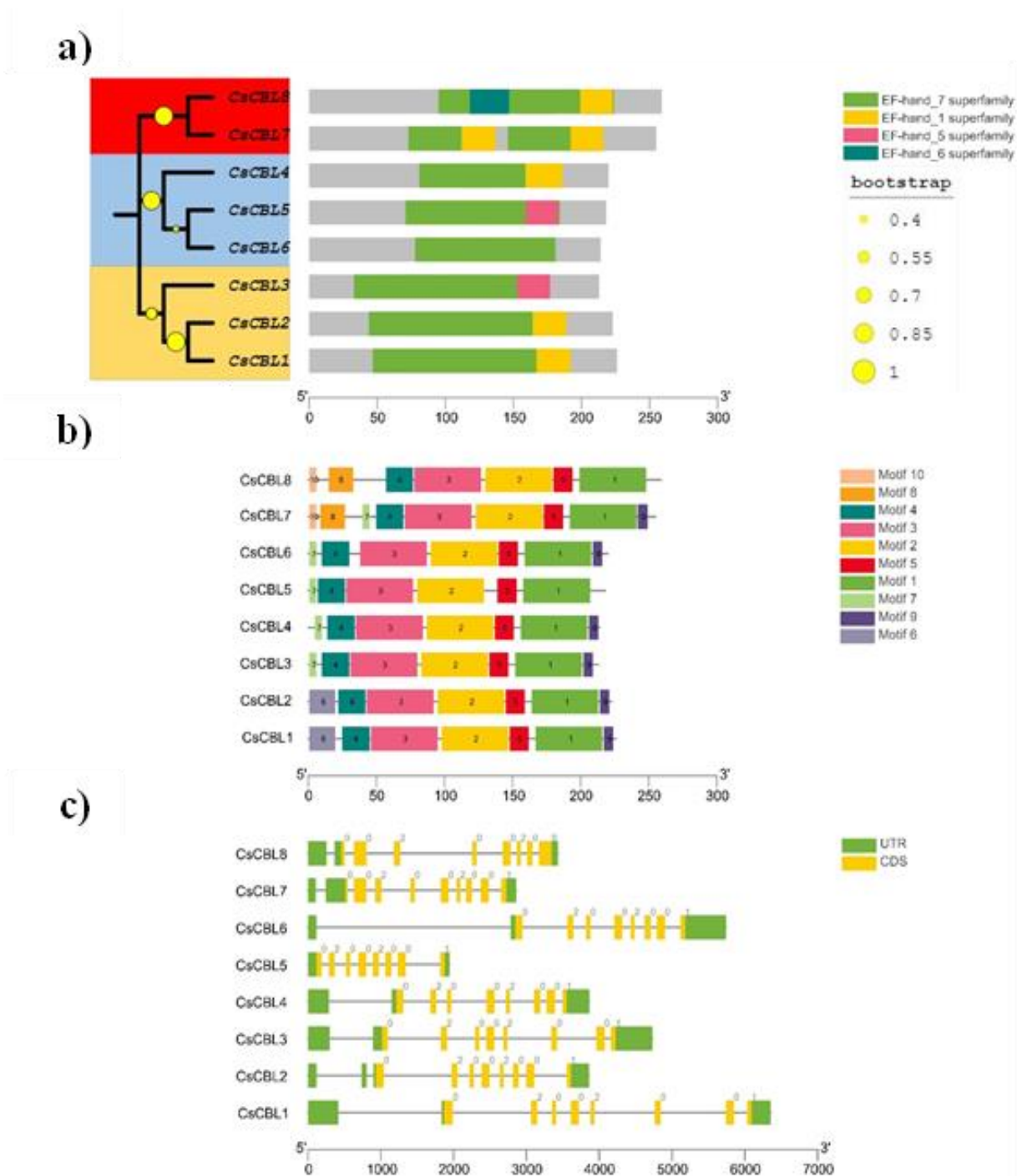
Figure 2. Schematic presentation of *CsCBL* gene structure, their domains and motifs organization. (a) Phylogenetic tree of eight predicted CsCBL as well as their EF-hands domain arrangement. (b) CsCBL motif distribution are shown in different colored boxes. (c) Gene structure of *CsCBL* gene family, and intron and exon organization are pictured. 5 and 3' untranslated -regions are marked by green boxes; the CDS is marked by yellow boxes; black lines denote introns. The bottom scale measures protein length.

Phylogenetic, conservative domains and motifs analysis of the CsCBLs

Protein sequence alignment was used to construct a phylogenetic tree, which was used to investigate the evolutionary relation between *A. thaliana, O. sativa, S. indicum* and *C. sinensis*. Three main groups (I to III) based on 37 CBL protein from different plant species could be seen in Figure 4. A total of seven, thirteen, and seventeen members were found in groups I, II, and III, respectively. In the first group, there were CsCBL7 and CsCBL8, in the second, CsCBL4, CsCBL5 and CsCBL6, and in the third, there were CsCBL1, CsCBL2, and CsCBL3. There was a disparity between the number of *C. sinensis* and *Arabidopsis* CBLs in each group, while the majority of the AtCBLs in group I clustered together. Molecular evolution is often examined at the gene or family level.

Protein families are sets of proteins that share at least 50% of their amino acid sequences. However, there are currently no models that can infer the evolution of gene families in order to estimate the ancestral state. Analysis of phylogenetic relationships and evolutionary events can be accomplished by using phylogenetic analysis.

Expression profiling of CsCBL

The RNA-seq data (TPM values) of *CsCBL* were obtained from GEO DataSets to examine the expression pattern in various sweet orange tissues, such as fruit, leaf, root, seed, flower, and stem, in order to comprehend the possible roles of the *CsCBL* family members. The expression profile of *CsCBL* genes in different bioprojects is shown in the column graph (Supplementary Figure S1). With the exception of *CsCBL4*, all seven *CsCBL* genes were found to be expressed in at least one tissue (Figure 5). The remarkable expression in most tissues, illustrating the critical functions of these CsCBL during sweet orange development and growth. The varied expression pattern of the *A. littoralus CBL* gene family in root and leaf tissues was proposed by s being associated with their functions in various biological processes and molecular functions. In this investigation, *CsCBL1* and *CsCBL8* were highly expressed in seeds, and *CsCBL5*, *CsCBL6* and *CsCBL7* was highly expressed in the stems.

Figure 3. Motif logo of CsCBL family were pictured by MEME program. MEME motifs are shown as stacks of letters. The X and Y axes reflect the motif's width and the number of bits in each letter.
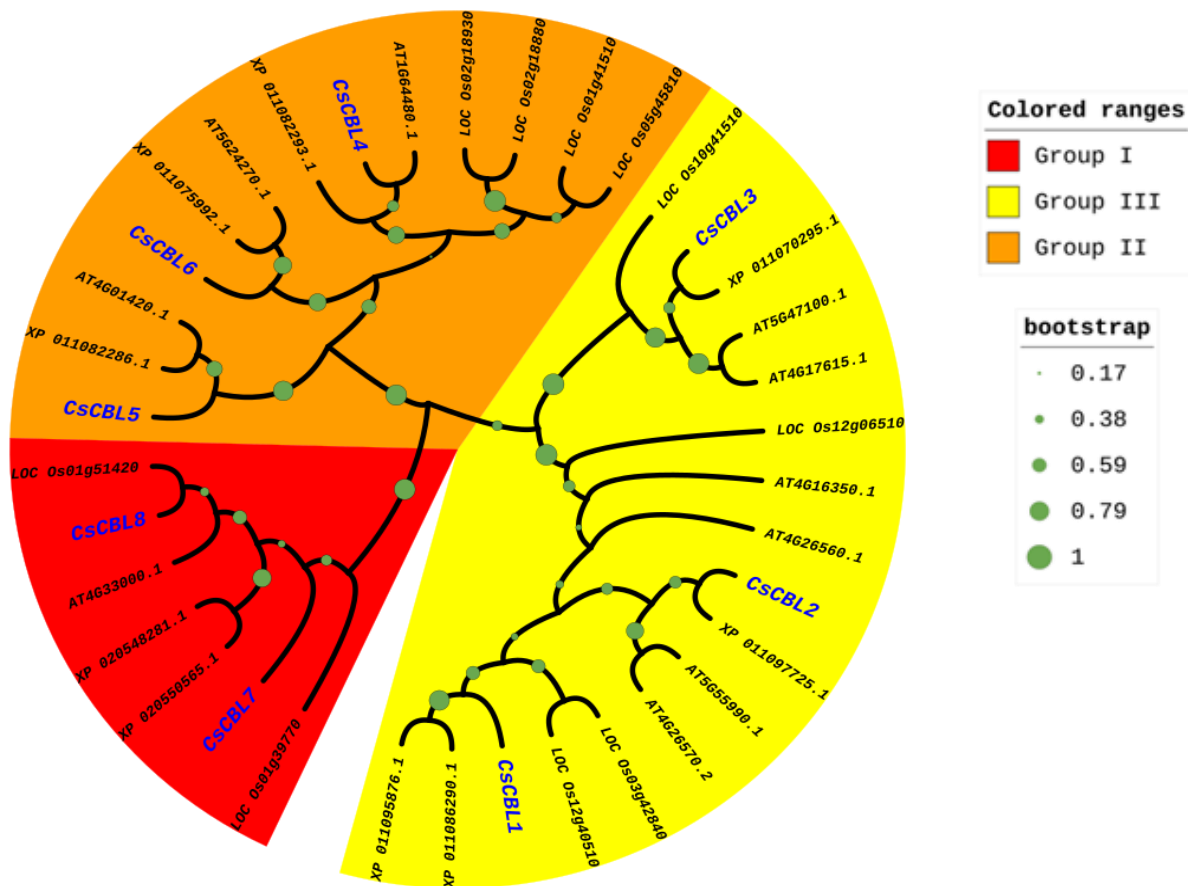
Figure 4. Phylogenetic tree of the CBL family based on an alignment of the proteins found in *Arabidopsis thaliana, Oryza sativa, sesamum indicum* and *C. sinensis*. Different CsCBL groups are represented by different-colored parts.

Tissue-specific expression was observed in the *CsCBL5* ortholog of the *Arabidopsis SOS3* gene (AT5G24270), with the highest levels of expression observed in stems and leaves and the lowest levels in fruit, root, seed, and flower tissues. It should be noted that, the expression of *CsCBL5*, in comparison to *CsCBL1*, *CsCBL2*, *CsCBL3*, and *CsCBL4*, was extremely low (TPM <5) in all of the specified tissues with the exception of leaves. This is likely due to the fact that *CsCBL5*, similar to *AtCBL4* (SOS3), serves regulatory functions that do not always necessitate high levels of expression.

Discussion

Abiotic and biotic stress responses are regulated by Ca^{2+} signaling in plants (Cheong et al., 2007; Ma et al., 2020). Furthermore, CDPKs, calmodulins, and calcineurin B-like proteins (CBLs) are all implicated in regulating Ca^{2+}-associated stress responses (Wang et al., 2016; Kudla et al., 2018). The availability of the *C. sinensis* genomic information (Wu et al., 2018) made it possible to conduct a genome-wide investigation of the *CsCBL* gene family. Due to their ubiquity in signaling cascades, Ca^{2+} sensor proteins have attracted increasing attention in recent years.

It was discovered in this study that there are eight CBL genes in *C. sinensis* that were found to be located on four chromosomes (Chr1, Chr02, Chr04 and Chr05 and one contig (chrNW-006257104.1).

Like *Arabidopsis*, CBLs in rice are encoded by a multigene family of at least ten members on chromosomes 1, 2, 3, 5, 10 and 12. In this study, the members of the *CsCBL* gene family share the same genetic characteristics.
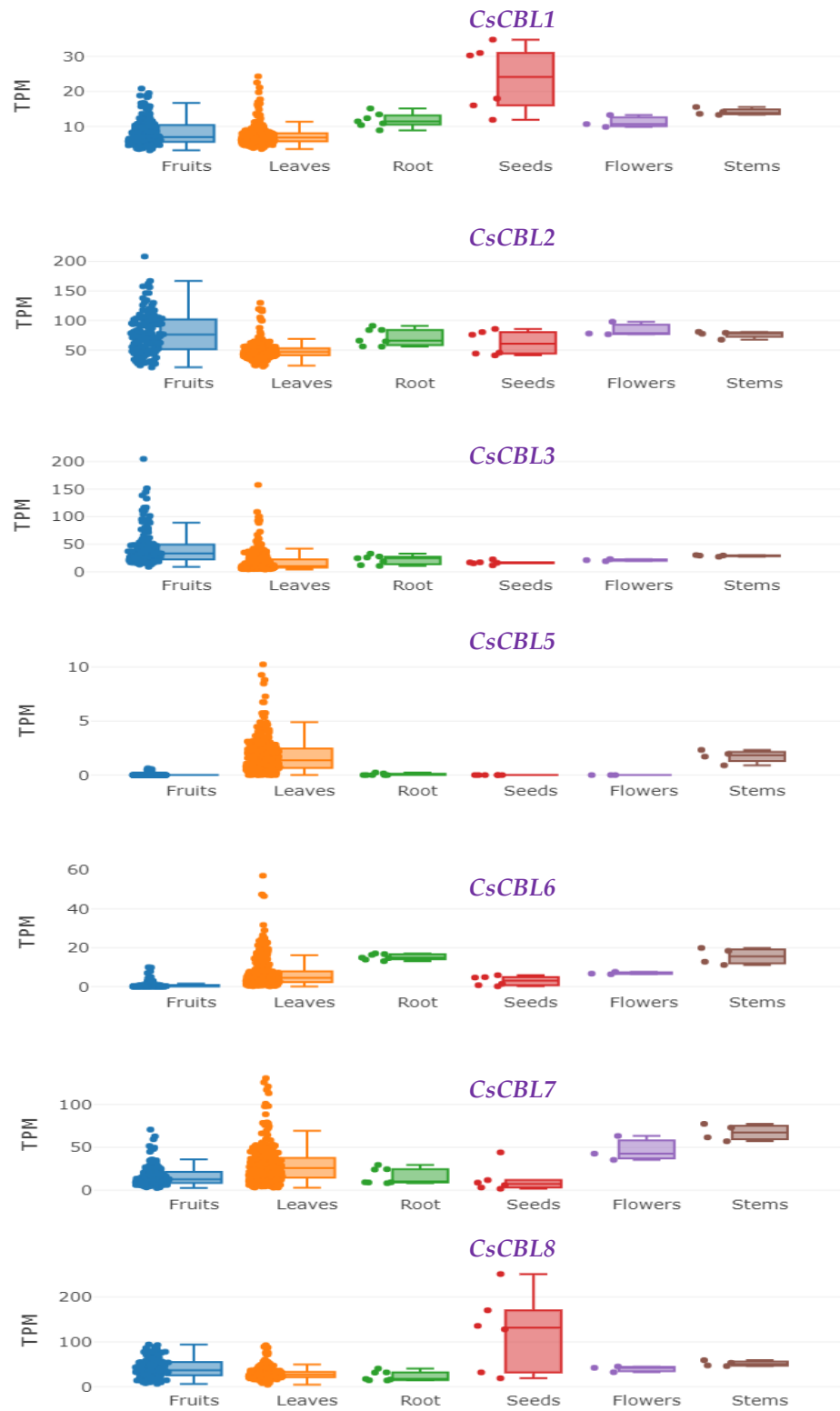
Figure 5. Meta-analysis of the *CBL* gene family's expression pattern in six distinct tissues of the sesame plant. The box plot displays expression levels in TPM (transcripts per million) format.

Among these *CsCBL* genes, polypeptides ranging from 24.42 to 29.67 kDa were predicted. The findings are comparable to those obtained from *A. thaliana* and *O. sativa*, where the majority of AtCBLs and OsCBLs had a size ranging from 23.5 to 25.9 kDa (Kolukisaoglu et al., 2004).

The coding regions of *CsCBL* gene are composed of either eight or nine exons. In other word, the majority of *CBL* genes in *C. sinensis* had seven or eight introns in their coding regions, which is very comparable to the number of introns found in *CBL* genes in *Arabidopsis*, rice, maize, canola, and eggplant (Kolukisaoglu et al., 2004; Gu et al., 2008; Zhang et al., 2014; Li et al., 2016; Zhang et al., 2016). On the other hand, the number of exons can affect the speed of gene induction and expression, and genes with fewer exons are expressed faster (Hashemipetroudi and Bakhshandeh, 2020; Hashemipetroudi et al., 2023; Yaghobi and Heidari, 2023). The junctions between their introns and exons follow the GT/AT rule (Mount, 1982). The length and position of the exons are quite consistent throughout the whole coding area. In addition to this, the majority of the CsCBLs have either three or two conserved EF hand domains in common with one another. The structures of the CsCBL proteins appear to be quite similar to one another, with each one containing four EF-hand domains. It's interesting to note that the linker region between the EF-hand domains appears to be unique to this family of calcium sensor proteins in terms of its constant size. Consequently, alterations in the size of CBL proteins are primarily associated with extension or reduction of the N- or C- terminus regions. As illustrated in Figure 2, the variation in protein size of CsCBLs can be attributed to the distinct dimensions of their amine-terminus (Gu et al., 2008). But some EF-hands are very different from the traditional EF-hand domain (Sathyanarayanan and Poovaiah, 2004; Sanchez-Barrena et al., 2005). AtCBL2 (Nagae et al., 2003) and SOS3, AtCBL4 (Sanchez-Barrena et al., 2005) also have sequences that differ significantly from the standard EF-hand sequence. These variations in the EF hand composition found in different CBLs can result in calcium ion binding affinities that are distinct from one another. Further, experimental investigations are required to determine whether or if these variations in calcium-binding affinity

contribute to the decoding of the many calcium signals that are produced in response to the myriad of environmental stimuli (Gu et al., 2008).

Myristoylation of proteins and palmitoylation of proteins are two critical processes that are necessary for protein trafficking, stability, and aggregation (Linder and Deschenes, 2007). According to Smotrys and Linder (2004), the process of protein myristoylation is initiated by the addition of myristic acid to the N-terminal glycine (Gly) amino acid, whereas the protein palmitoylation process is initiated by the addition of palmitic acid to the N-terminal cysteine (Cys) amino acid. The presence of myristoylated and palmitoylated regions in numerous CBLs may have facilitated the targeting of CBL-CIPK complex to membranes. *Arabidopsis*, rice, and other plants share many of these characteristics as well (Kolukisaoglu et al., 2004; Mohanta et al., 2015). It's possible that the highly comparable modes of action and/or interactions that these CBL family members share with their target CIPKs are reflected in the highly conserved structures of these CBL family members across diverse plant species (Mohanta et al., 2015). MGCV myristoylation sequence only found in CsCBL5 and CsCBL6 (Ishitani et al., 2000). Ancestral CBLs may have originated in this realm, according to a theory (Kleist et al., 2014). In the vast majority of the CsCBLs that have been investigated, the amino acid Gly located at the N-terminus is essential for protein myristoylation and is conserved at the second position. It was discovered that the seventh position in certain other CBL proteins contains a conserved form of the amino acid Gly at the N-terminal end of the protein. In a similar manner, the N-terminal Cys amino acid is necessary for protein palmitoylation, and it is conserved at the third position in CBL proteins. Certain CBLs do not have any N-terminal Cys amino acids in their structure. According to Blaskovic et al. (2013), protein palmitoylation is an ubiquitous alteration that is seen in membrane-bound proteins. This modification is also observed in transmembrane-spanning proteins that are generated in soluble ribosomes. In general, palmitoylation of proteins increases their ability to bind to membranes, which in turn changes their location and function. RasGTPase, Rho GTPase, and CDPKs are proteins that undergo palmitoylation (Mohanta et al., 2015). Except CsCBL3, CsCBL5, and

CsCBl6, all other CsCBLs lack a second N-terminal glycine. In addition, only CsCBL3, CsCBL5, and CsCBL6 have the N-terminal Gly amino acid at the second position in addition to having the cysteine amino acid at the N-terminus. Only myristoylation confers a modest affinity for membrane attachment, but palmitoylation and myristoylation confer extremely high affinity contacts (Martín and Busconi, 2001).

Among the CsCBLs, CsCBL5 had the highest aliphatic index and CsCBL7 the lowest. Alainine, valine, isoleucine, and leucine all include aliphatic side chains, and the aliphatic index of these amino acids has been found to be correlated with their respective proteins' thermostability (Ikai, 1980). Enzymes that have a better thermostability could be used in higher reaction temperatures, which would result in an accelerated reaction rate (by decrease of diffusional limitations). According to Sharafi et al. (2017), the half-life of thermostable enzymes is often longer than that of thermolabile enzymes. Stable enzymes are of particular relevance since they can be used in biocatalysis for longer periods of time (Sharafi et al., 2017).

37 CBLs were divided into three main groups by the phylogenetic tree produced from the protein sequence alignment of four species including *A. thaliana*, *O. sativa*, *S. indicum* and *C. sinensis*. The fact that members of the group shared comparable protein sequence lengths, motifs, and exon–intron architectures hints at a strong connection between them. Therefore, it is possible that CsCBLs and their homolog AtCBLs in the same branch play similar roles in plant-microbe interactions as well as in the responses to abiotic stress. According to Mohanta et al (2015), the sequence similarity of two genes indicates structural similarity, and structural similarity indicates functional similarity. As a result, it is possible that AtCBL1 and OsCBL1 perform similar roles. The strong similarities that exist between the CBL gene sequences suggest that they originated very recently as a result of gene duplication and may have functions that overlap with one another or are comparable to one another. Segmental duplication and tandem duplication are critical in the generation of additional members during the evolution of a gene family (Cannon et al., 2004). Due to this, the likely mechanism of *CsCBL* gene family evolution was examined by looking at both segmental and tandem duplication events. New functions led to evolution of paralogous genes, which are most likely to have a role in adaptation. In evolutionary biology, gene duplication and diversification are regarded to be the most significant processes. If a gene is duplicated, its second copy has less selective constraints and can evolve to have a little altered function while the original copy retains its original function.

In the majority of the tissues tested in the meta-analysis, the expression of a certain gene was extremely low. This is most likely because these genes perform regulatory roles that do not always demand high levels of expression. According to Liu et al. (2000), the low expression levels of *AtSOS2* and *AtSOS3* are related with their regulatory function in primary signal transduction. *CBL* proteins may be able to perform similar roles in distinct pathways due to their tissue-specific expression patterns. For example, according to (Guo et al., 2001), *AtCBL4* is expressed specifically in roots, whereas *AtCBL10* is expressed exclusively in shoot tissues (Kim et al., 2007).

Conclusion

The current study discovered eight CBL genes in the *C. sinensis* genome, which were classified into three groups based on phylogenetic analysis. Examining gene structure and expression patterns of gene families could provide useful information about their activity. The utilization of RNA-seq data in the meta-analysis of *CsCBL* gene expression indicated that the expression pattern profiles of *CsCBL* genes in the five tissues under investigation—roots, leaves, fruits, and stems—were discrete. The result of this study offers a unique opportunity to investigate calcium-related sensor in *C. sinensis*. It suggests that *CsCBL* can be utilized as a tool to examine the involvement of calcium ions (Ca^{2+}) in stress signaling, as well as to explore the various interaction of CBL and CIPK proteins in CBL-CIPK signaling network. More research into the expression of these genes at different biotic and abiotic sresses is needed to learn more about *CsCBL* gene regulation in general, and how these genes are controlled spatially and temporally in particular.

Supplementary Materials

The supplementary material for this article can be found online at: https://www.jpmb-gabit.ir/article_709799.html

Supplementary Table S1. *CsCBL* orthologue genes in *Arabidopsis*.

Supplementary Figure S1. The expression profile of *CsCBL* genes in different bioprojects is shown in the column graph.

Author Contributions

Conceptualization, S.H.H.; software, S.H.H. and M.A.; formal analysis, S.H.H. and M.A.; investigation, F.S. and H.G.; data curation, S.H.H.; writing—original draft preparation, S.H.H. and H.G.; writing—review and editing, S.H.H.; All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

Arab, M., Najafi Zarrini, H., Nematzadeh, G., Heidari, P., Hashemipetroudi, S.H., and Kuhlmann, M. (2023). analysis of calcium sensor families, CBL and CIPK, in *Aeluropus littoralis* and their expression profile in response to salinity. *Genes* 14(3): 753.

Blaskovic, S., Blanc, M., and van der Goot, F.G. (2013). What does palmitoylation do to membrane proteins? *FEBS J* 280(12): 2766-2774.

Cannon, S.B., Mitra, A., Baumgarten, A., Young, N.D., and May, G. (2004). The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. *BMC Plant Biol* 4(1): 10. doi: 10.1186/1471-2229-4-10.

Cheong, Y.H., Pandey, G.K., Grant, J.J., Batistic, O., Li, L., Kim, B.G., Lee, S.C., Kudla, J., and Luan, S. (2007). Two calcineurin B-like calcium sensors, interacting with protein kinase *CIPK23*, regulate leaf transpiration and root potassium uptake in *Arabidopsis*. *Plant J* 52(2): 223-239. doi: 10.1111/j.1365-313X.2007.03236.x.

Chow, C.N., Zheng, H.Q., Wu, N.Y., Chien, C.H., Huang, H.D., Lee, T.Y., Chiang-Hsieh, Y.F., Hou, P.F., Yang, T.Y., and Chang, W.C. (2016). PlantPAN 2.0: an update of plant promoter analysis navigator for reconstructing transcriptional regulatory networks in plants. *Nucleic Acids Res* 44(D1): D1154-1160. doi: 10.1093/nar/gkv1035.

Gao, P., Zhao, P.-M., Wang, J., Wang, H.-Y., Du, X.-M., Wang, G.-L., and Xia, G.-X. (2008). Co-expression and preferential interaction between two calcineurin B-like proteins and a CBL-interacting protein kinase from cotton. *Plant Physiol Biochem* 46(10): 935-940.

Gu, Z., Ma, B., Jiang, Y., Chen, Z., Su, X., and Zhang, H. (2008). Expression analysis of the calcineurin B-like gene family in rice (*Oryza sativa* L.) under environmental stresses. *Gene* 415(1-2): 1-12. doi: 10.1016/j.gene.2008.02.011.

Guo, Y., Halfter, U., Ishitani, M., and Zhu, J.-K. (2001). Molecular characterization of functional domains in the protein kinase SOS2 that is required for plant salt tolerance. *Plant Cell* 13(6): 1383-1400.

Hashemipetroudi, S., and Bakhshandeh, E. (2020). Expression analysis of *SiSOD* gene family during *Sesamum indicum* L. seed germination under various abiotic stresses. *J Plant Mol Breed* 8(2): 50-60.

Hashemipetroudi, S.H., Arab, M., Heidari, P., and Kuhlmann, M. (2023). Genome-wide analysis of the laccase (LAC) gene family in *Aeluropus littoralis*: A focus on identification, evolution and expression patterns

in response to abiotic stresses and ABA treatment. *Front Plant Sci* 14: 1112354. doi: 10.3389/fpls.2023.1112354.

Ikai, A. (1980). Thermostability and aliphatic index of globular proteins. *J Biochem* 88(6): 1895-1898.

Ishitani, M., Liu, J., Halfter, U., Kim, C.S., Shi, W., and Zhu, J.K. (2000). SOS3 function in plant salt tolerance requires N-myristoylation and calcium binding. *Plant Cell* 12(9): 1667-1678. doi: 10.1105/tpc.12.9.1667.

Kanchiswamy, C.N., Mohanta, T.K., Capuzzo, A., Occhipinti, A., Verrillo, F., Maffei, M.E., and Malnoy, M. (2013). Differential expression of CPKs and cytosolic Ca 2+ variation in resistant and susceptible apple cultivars (*Malus* x *domestica*) in response to the pathogen *Erwinia amylovora* and mechanical wounding. *BMC Genom* 14(1): 1-14.

Kim, B.G., Waadt, R., Cheong, Y.H., Pandey, G.K., Dominguez-Solis, J.R., Schultke, S., Lee, S.C., Kudla, J., and Luan, S. (2007). The calcium sensor CBL10 mediates salt tolerance by regulating ion homeostasis in *Arabidopsis*. *Plant J* 52(3): 473-484. doi: 10.1111/j.1365-313X.2007.03249.x.

Kleist, T.J., Spencley, A.L., and Luan, S. (2014). Comparative phylogenomics of the CBL-CIPK calcium-decoding network in the moss Physcomitrella, *Arabidopsis*, and other green lineages. *Front Plant Sci* 5: 187. doi: 10.3389/fpls.2014.00187.

Kolukisaoglu, Ü., Weinl, S., Blazevic, D., Batistic, O., and Kudla, J. (2004). Calcium sensors and their interacting protein kinases: genomics of the *Arabidopsis* and rice CBL-CIPK signaling networks. *Plant Physiol* 134(1): 43-58.

Kudla, J., Becker, D., Grill, E., Hedrich, R., Hippler, M., Kummer, U., Parniske, M., Romeis, T., and Schumacher, K. (2018). Advances and current challenges in calcium signaling. *New Phytol* 218(2): 414-431. doi: 10.1111/nph.14966.

Li, J., Jiang, M.M., Ren, L., Liu, Y., and Chen, H.Y. (2016). Identification and characterization of CBL and CIPK gene families in eggplant (*Solanum melongena* L.). *Mol Genet Genomics* 291(4): 1769-1781. doi: 10.1007/s00438-016-1218-8.

Li, L., Kim, B.G., Cheong, Y.H., Pandey, G.K., and Luan, S. (2006). A $Ca^{(2)+}$ signaling pathway regulates a $K^{(+)}$ channel for low K response in *Arabidopsis*. *Proc Natl Acad Sci USA* 103(33): 12625-12630. doi: 10.1073/pnas.0605129103.

Linder, M.E., and Deschenes, R.J. (2007). Palmitoylation: policing protein stability and traffic. *Nat Rev Mol Cell Biol* 8(1): 74-84. doi: 10.1038/nrm2084.

Liu, J., Ishitani, M., Halfter, U., Kim, C.-S., and Zhu, J.-K. (2000). The *Arabidopsis thaliana* SOS2 gene encodes a protein kinase that is required for salt tolerance. *Proc Natl Acad Sci* 97(7): 3730-3734.

Liu, L.L., Ren, H.M., Chen, L.Q., Wang, Y., and Wu, W.H. (2013). A protein kinase, calcineurin B-like protein-interacting protein Kinase9, interacts with calcium sensor calcineurin B-like Protein3 and regulates potassium homeostasis under low-potassium stress in *Arabidopsis*. *Plant Physiol* 161(1): 266-277. doi: 10.1104/pp.112.206896.

Lu, T., Zhang, G., Sun, L., Wang, J., and Hao, F. (2017). Genome-wide identification of CBL family and expression analysis of CBLs in response to potassium deficiency in cotton. *PeerJ* 5: e3653. doi: 10.7717/peerj.3653.

Luan, S. (2009). The CBL-CIPK network in plant calcium signaling. *Trends Plant Sci* 14(1): 37-42. doi: 10.1016/j.tplants.2008.10.005.

Ma, X., Li, Q.H., Yu, Y.N., Qiao, Y.M., Haq, S.U., and Gong, Z.H. (2020). The CBL-CIPK Pathway in Plant Response to Stress Signals. *Int J Mol Sci* 21(16): 5668. doi: 10.3390/ijms21165668.

Mao, J., Manik, S.M., Shi, S., Chao, J., Jin, Y., Wang, Q., and Liu, H. (2016). Mechanisms and Physiological Roles of the CBL-CIPK Networking System in *Arabidopsis thaliana*. *Genes* 7(9): 62. doi: 10.3390/genes7090062.

Martín, M.L., and Busconi, L. (2001). A rice membrane-bound calcium-dependent protein kinase is activated in response to low temperature. *Plant Physiol* 125(3): 1442-1449.

Mohanta, T.K., Mohanta, N., Mohanta, Y.K., Parida, P., and Bae, H. (2015). Genome-wide identification of Calcineurin B-Like (CBL) gene family of plants reveals novel conserved motifs and evolutionary aspects in calcium signaling events. *BMC Plant Biol* 15(1): 189. doi: 10.1186/s12870-015-0543-0.

Mohanta, T.K., and Sinha, A.K. (2016). "Role of calcium-dependent protein kinases during abiotic stress tolerance," in *Abiotic stress response plants,* eds. N. Tuteja & S.S. Gill. (USA: Wiley), pp. 185-206.

Mount, S.M. (1982). A catalogue of splice junction sequences. *Nucleic Acids Res* 10(2): 459-472. doi: 10.1093/nar/10.2.459.

Nagae, M., Nozawa, A., Koizumi, N., Sano, H., Hashimoto, H., Sato, M., and Shimizu, T. (2003). The crystal structure of the novel calcium-binding protein AtCBL2 from *Arabidopsis thaliana*. *J Biol Chem* 278(43): 42240-42246. doi: 10.1074/jbc.M303630200.

Sanchez-Barrena, M.J., Martinez-Ripoll, M., Zhu, J.K., and Albert, A. (2005). The structure of the *Arabidopsis thaliana* SOS3: molecular mechanism of sensing calcium for salt stress response. *J Mol Biol* 345(5): 1253-1264. doi: 10.1016/j.jmb.2004.11.025.

Sanyal, S.K., Pandey, A., and Pandey, G.K. (2015). The CBL-CIPK signaling module in plants: a mechanistic perspective. *Physiol Plant* 155(2): 89-108. doi: 10.1111/ppl.12344.

Sarwat, M., Ahmad, P., Nabi, G., and Hu, X. (2013). $Ca^{(2+)}$ signals: the versatile decoders of environmental cues. *Crit Rev Biotechnol* 33(1): 97-109. doi: 10.3109/07388551.2012.672398.

Sathyanarayanan, P.V., and Poovaiah, B.W. (2004). Decoding Ca^{2+} signals in plants. *Crit Rev Plant Sci* 23(1): 1-11. doi: 10.1080/07352680490273310.

Sharafi, E., Dehestani, A., and Farmani, J. (2017). Bioinformatics evaluation of plant chlorophyllase, the key enzyme in chlorophyll degradation. *Appl Food Biotech* 4(4): 167-178.

Smotrys, J.E., and Linder, M.E. (2004). Palmitoylation of intracellular signaling proteins: regulation and function. *Annu Rev Biochem* 73(1): 559-587. doi: 10.1146/annurev.biochem.73.011303.073954.

Steinhorst, L., Mähs, A., Ischebeck, T., Zhang, C., Zhang, X., Arendt, S., Schültke, S., Heilmann, I., and Kudla, J. (2015). Vacuolar CBL-CIPK12 Ca^{2+} sensor kinase complexes are required for polarized pollen tube growth. *Curr Biol* 25(11): 1475-1482.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol* 30(12): 2725-2729. doi: 10.1093/molbev/mst197.

Tang, R.J., Wang, C., Li, K., and Luan, S. (2020). The CBL-CIPK calcium signaling network: unified paradigm from 20 years of discoveries. *Trends Plant Sci* 25(6): 604-617. doi: 10.1016/j.tplants.2020.01.009.

Thoday-Kennedy, E.L., Jacobs, A.K., and Roy, S.J. (2015). The role of the CBL–CIPK calcium signalling network in regulating ion transport in response to abiotic stress. *Plant Growth Regul* 76(1): 3-12.

Tuteja, N., and Mahajan, S. (2007). Calcium signaling network in plants: an overview. *Plant Signal Behav* 2(2): 79-85. doi: 10.4161/psb.2.2.4176.

Wang, J.P., Xu, Y.P., Munyampundu, J.P., Liu, T.Y., and Cai, X.Z. (2016). Calcium-dependent protein kinase (CDPK) and CDPK-related kinase (CRK) gene families in tomato: genome-wide identification and functional analyses in disease resistance. *Mol Genet Genomics* 291(2): 661-676. doi: 10.1007/s00438-015-1137-0.

Wang, Y., Tang, H., Debarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.H., Jin, H., Marler, B., Guo, H., Kissinger, J.C., and Paterson, A.H. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res* 40(7): e49. doi: 10.1093/nar/gkr1293.

Wu, G.A., Terol, J., Ibanez, V., Lopez-Garcia, A., Perez-Roman, E., Borreda, C., Domingo, C., Tadeo, F.R., Carbonell-Caballero, J., Alonso, R., Curk, F., Du, D., Ollitrault, P., Roose, M.L., Dopazo, J., Gmitter, F.G., Rokhsar, D.S., and Talon, M. (2018). Genomics of the origin and evolution of Citrus. *Nature* 554(7692): 311-316. doi: 10.1038/nature25447.

Xu, J., Li, H.D., Chen, L.Q., Wang, Y., Liu, L.L., He, L., and Wu, W.H. (2006). A protein kinase, interacting with two calcineurin B-like proteins, regulates K^+ transporter AKT1 in *Arabidopsis*. *Cell* 125(7): 1347-1360. doi: 10.1016/j.cell.2006.06.011.

Yaghobi, M., and Heidari, P. (2023). Genome-wide analysis of Aquaporin gene family in *Triticum turgidum* and its expression profile in response to salt stress. *Genes* 14(1): 202. doi: 10.3390/genes14010202.

Yin, X., Wang, Q., Chen, Q., Xiang, N., Yang, Y., and Yang, Y. (2017). Genome-wide identification and functional analysis of the Calcineurin B-like protein and Calcineurin B-like protein-Interacting Protein Kinase gene families in Turnip (*Brassica rapa* var. rapa). *Front Plant Sci* 8: 1191. doi: 10.3389/fpls.2017.01191.

Yu, Q., An, L., and Li, W. (2014). The CBL-CIPK network mediates different signaling pathways in plants. *Plant Cell Rep* 33(2): 203-214. doi: 10.1007/s00299-013-1507-1.

Zhang, C., Bian, M., Yu, H., Liu, Q., and Yang, Z. (2011). Identification of alkaline stress-responsive genes of CBL family in sweet sorghum (*Sorghum bicolor* L.). *Plant Physiol Biochem* 49(11): 1306-1312. doi: 10.1016/j.plaphy.2011.08.010.

Zhang, F., Li, L., Jiao, Z., Chen, Y., Liu, H., Chen, X., Fu, J., Wang, G., and Zheng, J. (2016). Characterization of the calcineurin B-Like (CBL) gene family in maize and functional analysis of ZmCBL9 under abscisic acid and abiotic stress treatments. *Plant Sci* 253: 118-129. doi: 10.1016/j.plantsci.2016.09.011.

Zhang, H., Yang, B., Liu, W.Z., Li, H., Wang, L., Wang, B., Deng, M., Liang, W., Deyholos, M.K., and Jiang, Y.Q. (2014). Identification and characterization of *CBL* and *CIPK* gene families in canola (*Brassica napus* L.). *BMC Plant Biol* 14(1): 8. doi: 10.1186/1471-2229-14-8.

Zhang, H., Yin, W., and Xia, X. (2008). Calcineurin B-Like family in *Populus*: comparative genome analysis and expression pattern under cold, drought and salt stress treatment. *Plant Growth Regul* 56(2): 129-140.

شناسایی و تجزیه و تحلیل جامع خانواده ژنی *CBL* در پرتقال (.*Citrus sinensis* L)

سیدحمیدرضا هاشمی‌پطرودی ⁎۱، حمیدرضا قربانی ۲، فیروزه سهروردی ۱، مژده عرب ۱،۳

۱. پژوهشکده ژنتیک و زیست فناوری کشاورزی طبرستان، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری ایران

۲. بخش تحقیقات علوم زراعی و باغی، مرکز تحقیقات و آموزش کشاورزی و منابع طبیعی استان مازندران، سازمان تحقیقات، آموزش و ترویج کشاورزی، ساری، ایران.

۳. گروه زیست فناوری مولکولی گیاهی، پژوهشگاه ملی مهندسی ژنتیک و زیست فناوری، تهران، ایران

مقاله پژوهشی

چکیده: پروتئین‌های شبه‌کلسینورین B (CBL)، به عنوان حسگرهای کلسیم، نقش کلیدی در پاسخ‌های گیاه به عوامل تنش‌زای غیرزیستی مختلف و در رشد و نمو از طریق تعامل با پروتئین کینازهای متقابل CBL (CIPKs) ایفا می‌کنند. با این حال، اطلاعات در مورد نقش و کارکرد CBLs در گیاه پرتقال محدود است. در این تحقیق ژنوم گیاه *Citrus sinensis* مورد بررسی قرار گرفت و هشت ژن CBL شناسایی گردید. خصوصیت دامنه، موقعیت و توزیع موتیف‌های حفاظت‌شده پروتئینی نشان داد که دامنه EF-hands به طور نسبتاً زیادی در همه هشت CsCBL شناسایی شده حفظ شده است. پروتئین‌های CsCBL در گروه CBL اسیدی طبقه‌بندی شده و در این تحقیق پنج جایگاه myristoylation و شش جایگاه palmitoylation پیش‌بینی گردید. هشت ژن *CsCBL* بر روی چهار کروموزوم Chr01، Chr02، Chr04 و Chr05 و یک کانتیگ (chrNW-006257104.1) توزیع شده است. دو ژن *CsCBL4* و *CsCBL5* بر روی کروموزوم Chr05 پیوستگی داشته، که احتمالاً از طریق تکرارهای پشت سر هم ایجاد شده‌اند. درخت فیلوژنی ۳۷ پروتئین CBL از گونه‌های مختلف گیاهی از جمله آرابیدوپسیس تالیانا، برنج، کنجد و پرتقال نشان داد که این‌CBLها از ارتباط نزدیکی برخوردارند. نتایج این مطالعه ویژگی‌های عملکردی خانواده ژن *CsCBL* را برجسته نموده و اطلاعات ذی‌قیمتی را برای تحقیقات آینده در مورد فعالیت‌های عملکردی این ژن‌ها ارائه دهد.

کلمات کلیدی: پرتقال، پروتئین شبه‌کلسینورین B،تنش، حسگر کلسیم، پیام‌رسانی،CBL، *Citrus sinensis*.