Basic Analysis of Calcium-dependent Protein Kinase Gene and Its Closely Related Gene Families in Solanum Pennellii Genome

Jinyan Shi
China Agricultural University
Xiangge Du (1943597916@qq.com)
China Agricultural University

Research Article

Keywords: Solanum pennellii, calcium-dependent protein kinases (CDPK), CDPK-related protein kinase (CRK), ATP-binding region, Ser/Thr protein kinase region, EF-hand calcium-binding region, bioinformatics analysis

Posted Date: June 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-658642/v1

Abstract

Calcium-dependent protein kinases (CDPK) are the main Ca2 + sensor involved in the regulation of plant growth and development and various stress responses. In this study, we identified 32 CDPK (SpCDPK) genes and 7 CDPK-related protein kinase (SpCRK) genes in the whole genome of *Solanum Pennellii*, which were unevenly distributed on 12 chromosomes. The SpCDPK and SpCRK proteins own ATP-binding region and Ser/Thr protein kinase region. However, the SpCDPK proteins had EF-hand calcium-binding region, but the SpCRK proteins lacked it. Phylogenetic analysis showed that the SpCDPK and SpCRK gene families in *Solanum Pennellii* could be divided into four subgroups, and the evolutionary relationship between *Solanum Pennellii* and Arabidopsis thaliana was closer. Further analysis revealed that the exon-intron structure and conserved motif of each subgroup were basically the same, but there were differences in cis-acting elements. In this study, we conducted a preliminary analysis of SpCDPK and SpCRK gene families in *Solanum Pennellii* to provide basic data for further exploration of its molecular mechanism.

Introduction

Calcium ion (Ca2+) is a ubiquitous second messenger in animal, plant and microbial cells. It combines extracellular stimuli with its unique intracellular responses to regulate endogenous processes. To date, it is remarkable that three major Ca2 + sensor families have been identified in plants, including calcium-dependent protein kinases (CDPKs), calcineurin B-like (CBL) protein kinases (CIPKs) as well as calmodulin (CaMs)/ CaM-like proteins (CMLs)[1,2]. The elementary structure of the CDPK family genes show highly conserved in different species[3]. The CDPK comprises four characteristic domains: a variable N-terminal domain (involving the myristoylation and palmitoylation sites), a protein kinase domain (involving Ser/Thr protein kinase region and protein kinase ATP-binding region), an auto-inhibitory domain (acting as a pseudosubstrate combined with kinase domain to inhibit activity), and C-terminal (containing EF-hand motifs for Ca2 + binding)[4]. Most CDPK proteins contain myristoylation site (MGXXXSK) that related to protein localization (membrane localization) at the N terminus, indicating the substrates regulated by CDPK are mostly located on the cell membrane[5]. The Ser/Thr protein kinase region and EF-hand calcium-binding domain are the key functional areas for the transmission of calcium signals from CDPK protein to the downstream[6]. The C-terminal of CDPK protein accommodates 4 EF-hand structures, which is similar to CaM protein. The combination between EF-hands and Ca2 + activates CDPK protein to transmit signals downstream[7]. The CDPK-related protein kinases (CRK) are another protein kinase with Ser/Thr protein kinase region and protein kinase ATP-binding region, and they lack EF-hand calcium-binding domain[7]. In recent years, studies on plants such as of Arabidopsis, and rice have shown that CDPK and CRK families are composed of multiple members. The 8 CRK and 34 CDPK were identified in the entire genome of Arabidopsis[8], 5 CRK and 29 CDPK in *Oryza sativa*[9,10], 41 CDPK in *Gossypium raimondii*[11], 40 CDPK in maize[12], 25 CDPK in *Brassica napus*[13], 7 CDPK in *Vitis vinifera*[14], 9 CRK and 30 CDPK genes in *Populus trichocarpa*[15], 6 CRK and 29 CDPK in tomato[16,17], 5 CRKs and 31 CDPK in *Capsicum annuum*[18], 7 CRK and 18 CDPK in *Cucumis melo* L.[19].

CDPK genes are closely related to plant growth and development, hormone regulation and stress signal transduction. AtCPK28 is highly expressed in vascular and meristem and is involved in the regulation of shoot elongation and secondary growth. The *cpk28* mutant has extremely short stems[20]. OsCDPK5 and OsCDPK13 genes are highly expressed in root cortical cells, which stimulated RBOHH (Respiratory burst oxidase homolog isoform)-mediated ROS (Reactive oxygen species) production and resulted in the formation of normal aerenchyma in rice roots under hypoxia conditions[21]. AtCPK2/6/20 promotes pollen tube growth by activating SLAH3 (Slow anion channel 3) and ALMT12/13/14 (Aluminum-activated malate transporter 12/13/14)[22]. ZmCPK1 is a negative regulator in the response to cold stress[23], but OsCPK9 is identified in positively regulating drought stress tolerance[24]. Inhibition of MtCDPK1 gene expression result in shortening of root length and hair length[25]. StCDPK1 regulates auxin levels by phosphorylating the auxin transport vector StPIN4 (Auxin transport carrier)[26]. AtCPK6 positively regulates ABA signaling and drought tolerance by phosphorylating ABA responsive element binding factors[27]. Severe dwarfism and chlorosis occurred in Arabidopsis *crk1* mutants after continuous exposure to light, suggesting that AtCRK1 may be involved in light-regulated plant growth and development[28]. SlCRK6 positively regulates the resistance of tomato to PST DC3000 (Pseudomonas syringae pv. Tomato DC3000) and *Sclerotinia sclerotiorum*[17]. CmCRK2 expression is up-regulated and down-regulated respectively under cold stress and after inoculation with powdery mildew[19]. SlCDPK5/6 expression is up-regulated under drought stress, but is down-regulated at low temperature[16].

Solanum Pennellii has excellent stress resistance. Introducing more wild resources into tomato breeding is an important means to improve tomato varieties at present[29]. However, little is known about the function of the CDPK and CRK gene families in *Solanum Pennellii*. In this study, 7 CRK (SpCRK) genes and 32 CDPK (spCDPK) genes of *Solanum Pennellii* genome were collected for bioinformatics analysis. These data provided a theoretical basis for mining the function of CDPK gene.

Results

Identification and biochemical characteristics of CDPK and CRK genes in *Solanum Pennellii*

In this study, 39 non-redundant sequences were collected in the whole genome of *Solanum pennellii*, including 32 SpCDPK and 7 SpCRK. All of the SpCRK and SpCDPK proteins were designated separately as SpCRK1-SpCRK7 and SpCDPK1-SpCDPK32 based on their accession locus. By analyzing the physicochemical properties of SpCDPK and SpCRK proteins, it was found that the length of open reading frame of the SpCDPK proteins ranged from 511–607 aa, the molecular weight ranged from 55.9–68.0 kDa and the isoelectric point ranged from 5.0–9.0. The length of open reading frame, molecular weight, isoelectric point of SpCRK protein ranged from 501–607 aa, 56.4–68.2 kDa, 6.0-9.3, respectively (Table 1).

The feature domain analysis showed that the SpCRK carried ATP-binding region and Ser/Thr protein kinase region but lacked the EF-hand calcium-binding domain. However, except for SpCDPK3, SpCDPK5, SpCDPK27 and SpCDPK31, the ATP-binding region, Ser/Thr protein kinase region and four EF-hand calcium-binding domains existed in the remaining SpCDPK. Subcellular localization analysis showed that these genes acted in different organelles (Table 1).

Table 1
Physicochemical characterization of SpCRK and SpCDPK gene family proteins in *Solanum Pennellii*

Gene name	Accession locus	ORF length (aa)	Protein kinase domain		The number of EF-Hand calcium-binding region	Mol Wt (kDa)	pi	N-myristoylation site	N-Palmitoylation site	Sub-cell localization[a]
			Protein kinases ATP-binding region	Ser/Thr protein kinases region						
SpCRK1	XP_015055812	574	+	+		64.3	8.7	+	+	cyto
SpCRK2	XP_015062453	588	+	+		65.9	9.0	+	+	chlo
SpCRK3	XP_015063635	589	+	+		66.1	7.7	+	+	E.R.
SpCRK4	XP_015067044	607	+	+		68.0	9.0	+	+	chlo
SpCRK5	XP_015070667	511	+	+		55.9	6.0		+	cyto
SpCRK6	XP_015071034	598	+	+		66.9	9.0	+	+	nucl
SpCRK7	XP_015076220	587	+	+		65.8	8.5	+	+	cyto
SpCDPK1	XP_015055203	579	+	+	4	63.5	5.0		+	chlo
SpCDPK2	XP_015055554	544	+	+	4	60.4	5.5	+	+	cyto
SpCDPK3	XP_015055614	525	+	+	3	59.6	5.9	+	+	chlo
SpCDPK4	XP_015055658	557	+	+	4	62.2	5.6		+	nucl
SpCDPK5	XP_015056683	533	+	+	3	59.7	6.0	+	+	cyto
SpCDPK6	XP_015057806	528	+	+	4	59.4	5.7	+	+	cyto
SpCDPK7	XP_015058701	578	+	+	4	64.8	5.5		+	chlo
SpCDPK8	XP_015058992	505	+	+	4	56.8	5.4		+	nucl
SpCDPK9	XP_015059608	535	+	+	4	60.0	5.6		+	cyto
SpCDPK10	XP_015061184	525	+	+	4	58.4	6.5	+	+	mito
SpCDPK11	XP_015061185	518	+	+	4	57.5	6.5	+	+	mito
SpCDPK12	XP_015061191	535	+	+	4	59.6	5.4	+	+	cyto
SpCDPK13	XP_015063064	539	+	+	4	60.9	5.5	+	+	cyto
SpCDPK14	XP_015064109	570	+	+	4	64.3	9.3		+	chlo
SpCDPK15	XP_015064388	526	+	+	4	60.1	6.1	+	+	cyto
SpCDPK16	XP_015070509	573	+	+	4	64.9	9.0	+	+	chlo
SpCDPK17	XP_015070510	565	+	+	4	64.0	9.0	+	+	chlo
SpCDPK18	XP_015070549	582	+	+	4	64.6	5,6		+	chlo
SpCDPK19	XP_015070664	553	+	+	4	62.8	6.2	+	+	cyto
SpCDPK20	XP_015071018	538	+	+	4	60.9	6.4		+	cyto
SpCDPK21	XP_015071199	598	+	+	4	67.5	5.3		+	cyto
SpCDPK22	XP_015073898	527	+	+	4	59.5	5.3	+	+	cyto
SpCDPK23	XP_015074249	581	+	+	4	64.6	5.5		+	chlo
SpCDPK24	XP_015074284	508	+	+	4	57.2	5.1		+	cyto
SpCDPK25	XP_015076006	503	+	+	4	56.5	5.0		+	cyto
SpCDPK26	XP_015077134	501	+	+	4	56.4	5.6		+	chlo
SpCDPK27	XP_015078365	533	+	+	3	60.0	7.0	+	+	pero

Gene name	Accession locus	ORF length (aa)	Protein kinase domain		The number of EF-Hand calcium-binding region	Mol Wt (kDa)	pi	N-myristoylation site	N-Palmitoylation site	Sub-cell localization[a]
			Protein kinases ATP-binding region	Ser/Thr protein kinases region						
SpCDPK28	XP_015079858	536	+	+	4	61.0	5.8	+	+	mito
SpCDPK29	XP_015082059	521	+	+	4	57.8	6.6	+	+	pero
SpCDPK30	XP_015085844	516		+	4	57.7	5.8	+	+	chlo
SpCDPK31	XP_015088080	529	+	+	3	59.6	6.0		+	cyto
SpCDPK32	XP_015089561	607	+	+	4	68.2	5.4		+	chlo

[a] Nucl, nuclear; ER, endoplasmic reticulum; Mito, mitochondria; Cyto, cytosol; Chlo, chloroplast; Pero, Pero, peroxisomes[a] Nucl, nuclear; ER, endoplasmic reticulum; Mito, mitochondria; Cyto, cytosol; Chlo, chloroplast; Pero, Pero, peroxisomes

Chromosomal localization of SpCRK and SpCDPK genes

The chromosomal localization of SpCRK and SpCDPK genes were shown in Figure 2. The number of genes on chromosome 1 to chromosome 12 were 7, 3, 6, 3, 2, 2, 1, 1, 1, 6, 4 and 3, respectively. Chromosome 7, 8 and 9 involved only one SpCDPK gene. Chromosome 1 possessed the largest number of genes, with 7 genes (5 SpCDPK and 2 SpCRK). Chromosomes 2 and 3 each had 2 SpCRK genes, followed by chromosomes 2, 5 and 10 each had one SpCRK gene.

Protein sequence comparison of SpCRKs and SpCDPKs

In order to investigate the sequence characteristics of SpCRK and SpCDPK proteins, we performed multiple sequence alignments on the protein sequences of SpCRK and SpCDPK (Figure 2). As shown in Figure 2a, SpCRK and SpCDPK proteins contained protein kinase domains. Two highly conserved domains, the protein kinases ATP-binding region (LGxGxFGxTxCGxACKxIxK) and the Ser/Thr protein kinases region (VxHDRLKPENFLx), existed in the protein kinase domain.

The amino acid sequence analysis of the ATP-binding region of SpCDPK and SpCRK proteins was found that, compared with the SpCDPK proteins, the first amino acid in the ATP-binding region of SpCRK genes were mutated from L (leucine) to I (isoleucine)/V (Valine). And the F (Phenylalanine) mutation occurred in SpCDPK15, T (Threonine) mutation in SpCRK3 and SpCDPK15, two C (cysteine) mutations in SpCRK1-4, SpCRK6, SpCRK7, SpCDPK14, SpCDPK16, and SpCDPK17, only one C mutation in SpCDPK3 and SpCDPK30. Obviously, only the fourth G (glycines) in the ATP-binding region was mutated, and this phenomenon occurred in 10 genes. The last amino acid in the ATP-binding region, K (lysine), was mutated in SpCDPK22 and SpCDPK30. There were also a few proteins with amino acid mutations in the Ser/Thr protein kinases region. Two amino acid mutations existed in the Ser/Thr protein kinases region of SpCRK5, SpCDPK14, SpCDPK16 SpCDPK17, and SpCDPK32, while the Ser/Thr protein kinases region of the other genes were intact. Therefore, in the SpCRK and SpCDPK families, the Ser/Thr protein kinases region was more conserved than the ATP-binding region.

As shown in Figure 2b, all SpCDPKs contained the EF-hand calcium-binding domain (DxD/NxGxE), while SpCRKs lacked it. Significantly, the second D (Asparticacid) was mutated to G in the first EF-hand calcium-binding region of SpCDPK3, SpCDPK5, and SpCDPK27, which resulted in the deletion of an EF-hand calcium-binding region. It was also found that the second EF-hand calcium-binding region of SpDPK31 was incorrect, although it was consistent with the characteristic structure of the EF-hand region.

Phylogentic analysis of SpCRKs and SpCDPKs

The phylogenetic relationship about CDPK and CRK family members of *Solanum Pennellii* (32 SpCDPK and 7 SpCRK), Arabidopsis (34 CPK and 8 CRK) and rice (29 CPK and 5 CRK) were constructed via using the neighbor-joining method of MEGA6.0 (Figure 3). As Figure 3 showed, CDPK and CRK gene families were divided into four groups (I, II, III, and IV). The size of four groups were similar. The 10 CDPK in *Solanum Pennellii*, 13 CPK in Arabidopsis and 8 CPK in rice were put into group I. The group II included 13 SpCDPK, 10 AtCPK and 11 OsCPK. The group III comprised 6 *Solanum Pennellii* CDPK, 8 Arabidopsis CPK, and 8 rice CPK. The 20 CRK proteins from the three species were put into group IV. In addition to CRK, 3 CDPK (SpCDPK14, SpCDPK16 and SpCDPK17) of *Solanum Pennellii*, 2 CPK (AtCPK16, AtCPK18 and AtCPK28) of Arabidopsis and 3 CPK

(OsCPK4, OsCPK5 and OsCPK18) of rice existed in group IV. This phenomenon was not discovered in the other group. The dendrogram showed that the 39 proteins of Solanum Pennellii were generally closer to the proteins of Arabidopsis than rice, which indicated that they were evolutionarily more closely related.

Genetic structure analysis of SpCRK and SpCDPK genes

The genetic structure analysis of SpCRK and SpCDPK were carried out, and the results were shown in Figure 4. There were 4 groups in SpCRK and SpCDPK gene families, which were consistent with the respective corresponding phylogenetic relationships in Figure 3. As Figure 4 showed, genes within the same groups exhibited similar exon-intron organizations. The SpCDPK genes of the group I possessed 8-9 exons and 7-8 introns; group II genes included 7-8 exons and 6-7 introns; group III genes contained 7-9 exons and 6-8 introns. In particular, all SpCRK were part of the group IV. Except for SpCRK5 (involved 4 exons and 3 introns), all SpCRK introns and exons were 10 and 11, respectively. But SpCDPK of group IV contained 11 introns and 12 exons (Figure 4). These results showed that group IV exhibited more introns and exons than the other three groups, indicating that the gene structure of Group IV was more complex.

Conserved motif analysis of SpCRK and SpCDPK proteins

The results of motif analysis showed that the SpCRK and SpCDPK family proteins contained obvious structural characteristics. Motifs 9 and 10 were labeled as protein kinases ATP-binding region, motif 3 as Ser/Thr protein kinases region, and motifs 5, 6, 7 and 8 as EF-hand calcium-binding domain (Figure 6). As Figure 5 showed, interestingly, through analysis MEME online website, except for SpCRK5, the other 6 SpCRK included motifs 5, 6, 7 and 8; SpCDPK3, SpCDPK5 and SpCDPK27 included motif 5; and SpCDPK31 included motif 6. But these EF-hand calcium-binding region were not retrieved in Prosite (Table 1). It may be that these motifs were found only as recurring sequences, but they could not constitute EF-hand calcium-binding region structure. The SpCDPK of group I and II contained motifs 1-15, except for SpCDPK25 (Group II). The SpCDPK of Group III showed motifs 1-12 and 15, besides SpCDPK31. All proteins of group IV contained motifs 1-4, 10-11 and 15. These results indicated that all the identified proteins had typical ATP-binding region, Ser/Thr protein kinases region and EF-hand calcium-binding region, and each subgroup had similar motifs, which further supported the phylogenetic classification of SpCRK and SpCDPK families.

Promoter region analysis of SpCDPK and SpCRK genes

The cis elements in the upstream promoter region of SpCRK and SpCDPK genes were analyzed by PlantCARE. The results showed that the promoter region of SpCRK and SpCDPK genes contained 627 cis-acting elements, which can be divided into five types: hormone-related elements, growth-related elements, stress-related elements, secondary metabolite-related element and plant protein metabolism-related element (Figure 7). The hormone-related elements included auxin responsive element, gibberellin responsive element, salicylic acid responsive element, abscisic acid responsive element and MeJA responsive element. The growth-related elements included circadian control responsive element, palisade mesophyll cell differentiation responsive element, endosperm expression response element, seed-specific regulation responsive element, light responsive element and meristem expresses related elements. The stress-related elements included low temperature responsive element, drought responsive element, anoxic induction element, anaerobic induction element. The secondary metabolite-related element and plant protein metabolism-related element was flavonoid biosynthetic genes regulatory elements and zein metabolism regulatory elements, respectively. Among all the elements, the number of light responsive elements was the largest, with 356. Obviously, the number and distribution of cis-acting elements of SpCDPK10 and SpCDPK11 genes were basically the same, and SpCDPK16 and SpCDPK17 also showed similar phenomenon. All elements of SpCDPK20 were located between 900bp and 2000bp upstream sequence. The cis-acting elements of SpCDPK27 were mainly distributed at both ends of the promoter region (Figure 7). These results suggest that SpCRK and SpCDPK genes play an important role not only in the growth and development of *Solanum Pennellii*, but also in the response to biotic and abiotic stresses.

Discussion

In this study, a total of 7 CRK and 32 CDPK genes were retrieved from the whole genomic data of *Solanum Pennellii*, which was similar to the number of CRK and CDPK genes in Arabidopsis (5 CRK and 29 CDPK) and rice (8 CRK and 34 CDPK)[30-33]. The *Solanum lycopersicum* genome, however, contained only the 29 CDPK genes and lacked the CRK genes[34]. The 29 SLCDPK genes encoded peptides with a range of 429-598 aa, and the predicted molecular weights of the proteins were 48.06-67.55 kDa, respectively. The isoelectric point of most SlCDPKs showed alkalinity[34]. The length of open reading frame, molecular weight of 32 SpCDPK and 7 SpCRK ranged from 501-607 aa, 55.9-68.2 kDa, respectively (Table 1). These results showed that the SpCDPK and SpCRK proteins were longer and had higher molecular weights than those of

Solanum lycopersicum. Interestingly, 29 of the 39 genes (32 SpCDPK and 7 SpCRK) were acidic, the opposite of what was found in tomatoes (Table 1)[34]. These results suggested that tomatoes evolved to eliminate redundant structures and creating a simpler genetic structure for survival. The myristoylation or palmitylation sites played a role in the binding of proteins to membranes[35]. The SpCDPKs and SpCRKs contained myristoylation or palmitylation sites at the N terminal, indicating that they might be located on organelle membranes (Table 1).

The myristoylation site caused the irreversible loose structure between protein and membrane, while the palmitylation site caused the reversible stable structure between protein and membrane[36]. The subcellular localization of SpCDPKs and SpCRKs were diverse, including nuclear, endoplasmic reticulum, mitochondria, cytosol, chloroplast and peroxisomes localization, suggesting they own varied function (Table 1)[37,38]. It was also found that the localization of CDPKs and CRKs proteins in cells was also changed after the mutation of the palmitoylation site or the myrimyoylation site[39,40].

The SpCDPK and SpCRK family members were unevenly distributed on 12 chromosomes of wild tomato, and most of the family members were located at the front or end of chromosomes (Figure 1). This was also found in the studies of the CDPK gene family of Solanum lycopersicum[34]. By protein sequence alignment, SpCDPK gene existed three domains: ATP-binding region, Ser/Thr protein kinases region and EF-hand calcium-binding domain. The C-terminal EF-hand calcium-binding domain of SpCRK had been degenerated, and other structures were similar to SpCDPK (Figure 2). This result was also found in conserved motif analysis (Figure 5 and Figure 6). Due to the degenerated EF-hand structure at the C-terminal of most CRKs, CRKs were not directly regulated by Ca2+. However, some CRKs contained CaM binding regions, the kinase activity of CRKs was under the coordinated control of Ca2+ and CaM[41,42].

The results of phylogenetic analysis showed that SpCDPK and SpCRK gene families were divided into four groups: Ⅰ-Ⅳ. The SpCDPK and SpCRK genes were closer to Arabidopsis in evolutionary relationship, while CDPK and CRK family members of monocotyledonous rice were significantly different from those of dicotyledonous plants in evolutionary relationship (Figure 3).

Studies have shown that CDPK and CRK genes were involved in the regulation of root, stem, leaf development, flowering, pollen germination, pollen tube growth and seed development[4, 21, 22, 28, 43-47], as well as in the regulation of gibberellin, auxin, salicylic acid, abscisic acid and methyl jasmonate biosynthesis[27,48-53]. In addition, they also played a role in biological and abiotic stress responses[54-64]. By analyzing the cis-action elements of 2000 bp promoters upstream of SpCDPK and SpCRK family members, we also found that there were a large number of elements related to plant growth and development, hormone induction and stress induction (Figure 7).

The results of this experiment laid a foundation for the functional analysis of SpCDPK and SpCRK genes, and provided some theoretical guidance for the further study of the function of SpCDPK and SpCRK genes in tomato stress resistance and the cultivation of good quality.

Materials And Methods

Identification and physicochemical characterization of SpCDPK and SpCRK

All CDPK and CRK protein sequences of Arabidopsis and rice were obtained by querying the TAIR database (The Arabidopsis Information Resource, http://www.arabidopsis.org/) and rice Database (http://rice.plantbiology.msu.edu/), respectively. The whole proteins and nucleotide sequences of Solanum Pennellii were searched from NCBI (https://www.ncbi.nlm.nih.gov/genome/).

Online tools ExPASy (http://web.expasy.org/myristoylator/)[65] and Wolf PSort (http://www.genscript.com/psort/wolf_psort.html)[66] and GPS-Palm[67] software were used to analyze the physicochemical properties, subcellular localization and palmitoylation site of SpCDPK and SpCRK gene families, respectively. The Prosite tool of ExPASy was used to retrieve the EF-hand calcium-binding domain and protein kinase domain. The online MapGene2Chrom program was used to map the position of genes on chromosomes (http://mg2c.iask.in/mg2c_v2.0/)[68].

Multi-sequence alignment and phylogenetic tree construction of SpCDPK and SpCRK protein

The SpCDPK and SpCRK protein sequences of *Solanum Pennellii* were aligned by MUSCLE program of MEGA6.0[69] with 70% conserved sites . Then, phylogenetic trees were constructed using the neighbor-joining method of MEGA6.0, in which bootstrap was set to 1000, based on sequence alignment results.

Structure analysis of protein and gene

The exons and introns of these genes was analyzed with Online website GSDS (http://gsds.cbi.pku.edu.cn/)[70,71]. The MEME suite (http://alternate.meme-suite.org/tools/meme)[72] was used to retrieve the motif of the SpCDPK and SpCRK proteins, and the motif number was set to 15. The cis-acting elements of 2 000 bp upstream promoter were analyzed by the Plant CARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/)[73].

Declarations

Acknowledgments

This work was supported by grants from the Beijing innovation team construction project (6514001). The funders had no role in study design, data collection, data analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceived and designed the experiments: J.S. and X.D. Performed the experiments: J.S. Analyzed the data: J.S. Wrote the paper: J.S. Provided guidance on the whole study: J.S and X.D.

Additional Information

The authors declare no competing interests.

References

1. Xu, W. W. & Chuang, W. *Calcium-dependent protein kinases in phytohormone signaling pathways.International Journal Molecular Sciences11,2346–2361*(*2017*).
2. Dodd, A. N. *et al. The language of calcium signaling.Annual Review of Plant Biology61,593–620*(*2010*).
3. Valmonte, G. R. *et al. Calcium-dependent protein kinases in plants*: *evolution*, *expression and function.Plant Cell Physiol. 55,551–569*(*2014*).
4. Delormel, T. Y. & Boudsocq, M. *Properties and functions of calcium-dependent protein kinases and their relatives in Arabidopsis thaliana.New Phytologist224,585–604*(*2019*).
5. Schulz, P. *et al. Calcium-dependent protein kinases*: *hubs in plant stress signaling and development.Plant Physiol. 163,523–530*(*2013*).
6. Ludwig, A. A. *et al. CDPK-mediated signalling pathways*: *specificity and cross-talk.J. Exp. Bot. 55,181–188*(*2004*).
7. Harmon, A. C. *et al. CDPKs-a kinase for every Ca2 + signal?Trends Plant Sci. 5,154-9*(*2000*).
8. Cheng, S. H. *et al. Calcium signaling through protein kinases. The Arabidopsis calcium-dependent protein kinase gene family.Plant Physiol. 129,469–485*(*2002*).
9. Asano, T. *et al. Genome-wide identification of the rice calcium-dependent protein kinase and its closely related kinase gene families*: *comprehensive analysis of the CDPKs gene family in rice.Plant Cell Physiol. 46,356–366*(*2005*).
10. Ye, S. *et al. Expression profile of calcium-dependent protein kinase (CDPKs) genes during the whole lifespan and under phytohormone treatment conditions in rice (Oryza sativa L. ssp. indica).Plant Mol. Biol. 70,311–325*(*2009*).
11. Liu, W. *et al. Genome-wide survey and expression analysis of calcium-dependent protein kinase in Gossypium raimondii.PLoS One9,e98189*(*2014*).
12. Kong, X. P. *et al. Genome-wide identification and expression analysis of calcium-dependent protein kinase in maize.BMC Genomics14,433*(*2013*).
13. Zhang, H. *et al., Identification, expression and interaction analyses of calcium-dependent protein kinase (CPK) genes in canola (Brassica napus L.).BMC Genom. 15,211*(*2014*).
14. Chen, F. *et al. The evolutionary history and diverse physiological roles of the grapevine calcium-dependent protein kinase gene family.PLoS One8,e80818*(*2013*).
15. Zuo, R. *et al. Genome-wide identification, classification, and expression analysis of CDPK and its closely related gene families in poplar (Populus trichocarpa).Mol. Biol. Rep. 4,2645–2662*(*2013*).
16. Hu, Z. *et al. Genome-wide identification and expression analysis of calcium-dependent protein kinase in tomato.Front. Plant Sci. 7,469*(*2016*).

17. Wang, J. P. *et al. 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 291,661–676*(*2016*).
18. Cai, H. *et al. Genome-wide identification and expression analysis of calcium-dependent protein kinase and its closely related kinase genes in Capsicum annuum.Front. Plant Sci. 6,737*(*2015*).
19. Zhang, H. *et al. Genome-wide identification and expression analysis of calciumdependent protein kinase and its related kinase gene families in melon (Cucumis melo L.).PLoS One12,e0176352*(*2017*).
20. Matschi, S. *et al. Function of calcium-dependent protein kinase CPK28 of Arabidopsis thalianain plant stem elongation and vascular development.The Plant Journal73,383–396*(*2013*).
21. Yamauchi *et al., An NADPH oxidase BOH functions in rice roots during lysigenous aerenchyma formation under oxygen-deficient conditions.The Plant Cell29,775–790*(*2017*).
22. Gutermuth, T. *et al. Tip-localized Ca2+ -permeable channels control pollen tube growth via kinase-dependent R- and S-type anion channel regulation.New Phytologist218,1089–1105*(*2018*).
23. Weckwerth, P. *et al. ZmCPK1, a calcium-independent kinase member of the Zea mays CDPK gene family, functions as a negative regulator in cold stress signalling,Plant Cell Environ. 38,544–558*(*2015*).
24. Wei, S. *et al. A rice calcium-dependent protein kinase OsCPK9 positively regulates drought stress tolerance and spikelet fertility.BMC Plant Biol. 14,133*(*2014*).
25. Ivshuta, S. *et al. RNA interference identifies a calcium-dependent protein kinase involved in Medicago truncatula root development.The Plant Cell17,2911–2921*(*2005*).
26. Santin, F. *et al. Solanum tuberosum StCDPK1 is regulated by miR390 at the posttranscriptional level and phosphorylates the auxin efflux carrier StPIN4 in vitro, a potential downstream target in potato development.Physiologia Plantarum159,244–261*(*2017*).
27. Zhang, H. F. *et al. Arabidopsis CPK6 positively regulates ABA signaling and drought tolerance through phosphorylating ABA-responsive element-binding factors.Journal of Experimental Botany71,188–203*(*2020*).
28. Baba, A. I. *et al. Functional analysis of the Arabidopsis thaliana CDPK-related kinase family: AtCRK1 regulates responses to continuous light.International Journal of Molecular Sciences19,1282*(*2018*).
29. Du, M. M. *et al. Current status and prospects on tomato molecular breeding-from gene cloning to cultivar improvement.Acta Hort Sin. 44,581–600*(*2017*).
30. Zhao, Y. *et al. Genome-wide identification and functional analyses of calmodulin genes in Solanaceous species.BMC Plant Biology13,70*(*2013*).
31. Mohanta, T. K. *et al. Genomics and evolutionary aspect of calcium signaling event in calmodulin and calmodulin-like proteins in plants.BMC Plant Biol. 17,38*(*2017*).
32. McCormack, E. & Braam, J. *Calmodulins and related potential calcium sensors of Arabidopsis.New Phytologist159,585–598*(*2003*).
33. Boonburapong, B. & Buaboocha, T. *Genome-wide identification and analyses of the rice calmodulin and related potential calcium sensor proteins.BMC Plant Biol. 7,4*(*2007*).
34. Munir, S. *et al. Genome-wide identification, characterization and expression analysis of calmodulin-like (CML) proteins in tomato (Solanum lycopersicum).Plant Physiol. Biochem. 102,167–179*(*2016*).
35. Rutschmann, F. *et al. LeCPKl, a calcium-dependent protein kinase from tomato, Plasma membrane targeting and biochemical characterization.Plant physiology129,156–168*(*2002*).
36. Resh, M. D. *Trafficking and signaling by fatty-acylated and prenylated proteins,Nature Chemical Biology2,584–590*(*2006*).
37. Boudsocq, M. & Sheen, J. *CDPKs in immune and stress signaling.Trends in Plant Science18,30–40*(*2013*).
38. Simeunovic, A. *et al. Know where your clients are: Subcellular localization and targets of calcium-dependent protein kinases.Journal of Experimental Botany67,3855–3872*(*2016*).
39. Stael, S. *et al. Protein N-acylation overrides differing targeting signals.Febs Letters585,517–522*(*2011*).
40. Mehlmer, N. *et al. The Ca2+-dependent protein kinase CPK3 is required for MAPK-independent salt-stress acclimation in Arabidopsis.The Plant Journal63,484–498*(*2010*).
41. Popescu, S. C. *et al.* Differential binding of calmodulin-related proteins to their targets revealed through high-density Arabidopsis protein microarrays. Proceedings of the National Academy of Sciences 104, 4730–4735 (*2007*).
42. Wang, Y. *et al. Characterization of a calmodulin-regulated Ca2+-dependent-protein-kinase-related protein kinase, AtCRK1, from Arabidopsis.Biochemical Journal383,73–81*(*2004*).
43. Rig, G. *et al. Inactivation of plasma membrane-localized CDPK related kinase 5 decelerates PIN2 exocytosis and root gravitropic response in Arabidopsis.The Plant Cell25,1592–1608*(*2013*).

44. Wen, K. X. *et al. OsCPK21 is required for pollen latestage development in rice.Journal of Plant Physiology*. *240,1–3*(*2019*).
45. Jiang, J. Z. *et al. Effects of OsCDPK1 on the structure and physicochemical properties of starch in developing rice seeds.International Journal of Molecular Sciences19,3247*(*2018*).
46. Yang, D. H. *et al. Silencing Nicotiana attenuata calcium-dependent protein kinases*, *CDPK4 and CDPK5*, *strongly up-regulates wound- and herbivory-induced jasmonic acid accumulations.Plant Physiology159,1591–1607*(*2012*).
47. Lee, S. S. *et al. Iteraction of NtCDPK1 calcium-dependent protein kinase with NtRpn3 regulatory subunit of the 26S proteasome in Nicotiana tabacum.The Plant Journal33,825–840*(*2003*).
48. Wei, S. Y. *et al. A rice calcium-dependent protein kinase OsCPK9 positively regulates drought stress tolerance and spikelet fertility.BMC Plant Biology14,1–13*(*2014*).
49. Matschi, S. *et al. The calcium-dependent protein kinase CPK28 regulates development by inducing growth phase-specific*, *spatially restricted alterations in jasmonic acid levels independent of defense responses in Arabidopsis.The Plant Cell27,591–606*(*2015*).
50. Chen, D. H. *et al. Calcium-dependent protein kinase CPK9 negatively functions in stomatal abscisic acid signaling by regulating ion channel activity in Arabidopsis.Plant Molecular Biology99,113–122*(*2019*).
51. Huang, S. J. *et al. A type III ACC synthase*, *ACS7*, *is involved in root gravitropism in Arabidopsis thaliana.Journal of Experimental Botany64,4343–4360*(*2013*).
52. Rietz, S. *et al. Roles of Arabidopsis patatin-related phospholipases a in root development are related to auxin responses and phosphate deficiency.Molecular Plant3,524–538*(*2010*).
53. Ishida, S. *et al. A tobacco calcium-dependent protein kinase*, *CDPK1*, *regulates the transcription factor repression of shoot growth in response to gibberellins.The Plant Cell20,3273–3288*(*2008*).
54. Bund, M. & Coca, M. *Enhancing blast disease resistance by overexpression of the calcium-dependent protein kinase OsCPK4 in rice.Plant Biotechnology Journal14,1357–1367*(*2016*).
55. Boller, T. & He, S. Y. *Innate immunity in plants*: *an arms race between pattern recognition receptors in plants and effectors in microbial pathogens.Science324,742–744*(*2009*).
56. Bund, M. & Coca, M. *Calcium-dependent protein kinase OsCPK10 mediates both drought tolerance and blast disease resistance in rice plants.Journal of Experimental Botany68,2963–2975*(*2017*).
57. Borkiewicz, L. *et al. Expression of maize calcium-dependent protein kinase (ZmCPK11) improves salt tolerance in transgenic Arabidopsis plants by regulating sodi- um and potassium homeostasis and stabilizing photosystem II.Physiologia Plantarum168,38–57*(*2020*).
58. Liu, N. *et al. Calcium-dependent protein kinase 5 associates with the truncated NLR protein TIR-NBS2 to contribute to exo70B1-mediated immunity.The Plant Cell29,746–759*(*2017*).
59. Boudsocq, M. *et al. Differential innate immune signalling via Ca2 + sensor protein kinases.Nature464,418–422*(*2010*).
60. Wang, J. Y. *et al. The kinase OsCPK4 regulates a buffering mechanism that fine-tunes innate immunity.Plant Physiology176,1835–1849*(*2018*).
61. Liu, Y. *et al. The calcium-dependent kinase OsCPK24 functions in cold stress responses in rice.Journal of Integrative Plant Biology60,173–188*(*2018*).
62. Zhang, H. L. *et al. The Arabidopsis Ca2+-dependent protein kinase CPK12 is involved in plant response to salt stress.International Journal of Molecular Sciences19,1–17*(*2018*).
63. Almadanim, M. C. *et al. Rice calcium-dependent protein kinase OsCPK17 targets plasma membrane intrinsic protein and sucrose-phosphate synthase and is required for a proper cold stress response.Plant Cell and Environment40,1197–1213*(*2017*).
64. Xie, K. B. *et al. Direct phosphorylation and activation of a mitogen-activated protein kinase by a calcium-dependent protein kinase in rice*.*The Plant Cell26,3077–3089*(*2014*).
65. Séverine, D. *et al. Expasy*, *the Swiss Bioinformatics Resource Portal*, *as designed by its users.Nucleic Acids Research*(*2021*)
66. Horton, P. *et al. WoLF PSORT*: *protein localization predictor.Biochimie*. *35,W585-587*(*2007*).
67. Ning, W. S. *et al. GPS-Palm*: *a deep learning-based graphic presentation system for the prediction of S-palmitoylation sites in proteins.Briefings in Bioinformatics*(*2020*).
68. Chao, J. T. *et al. MapGene2Chrom*, *a tool to draw gene physical map based on Perl and SVG languages.Hereditas37,91–97*(*2015*).
69. Tamura, K. *et al. MEGA6*: *molecular evolutionary genetics analysis version 6.0.Mol. Biol. Evol. 30,2725–2729*(*2013*).
70. Hu, B. *et al. GSDS 2.0*: *an upgraded gene feature visualization server.Bioinformatics31,1296–1297*(*2015*).
71. Qiao, L. *et al. Characterization and expression patterns of auxin response factors in Wheat.Front. Plant Sci. 9,1395*(*2018*).
72. Brown, P. *et al. MEME-LaB*: *motif analysis in clusters.Bioinformatics29,1696–1697*(*2013*).

73. Lescot, M. *et al. PlantCARE: a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences.Nucleic. Acids. Res. 30,325-7(2002*).

Figures

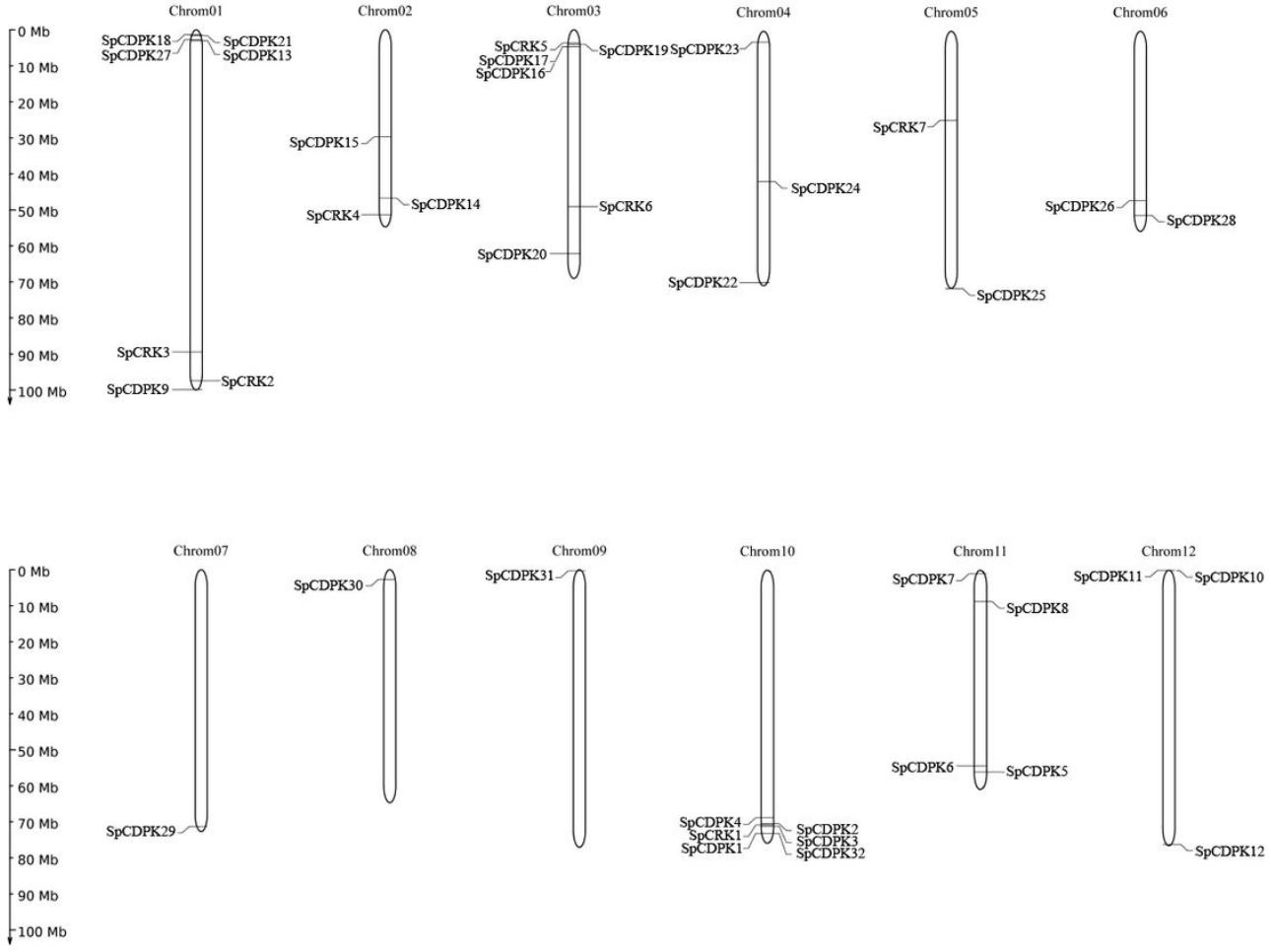

Figure 1

Chromosomal localization of SpCRK and SpCDPK genes. The scale is in megabass (Mb).

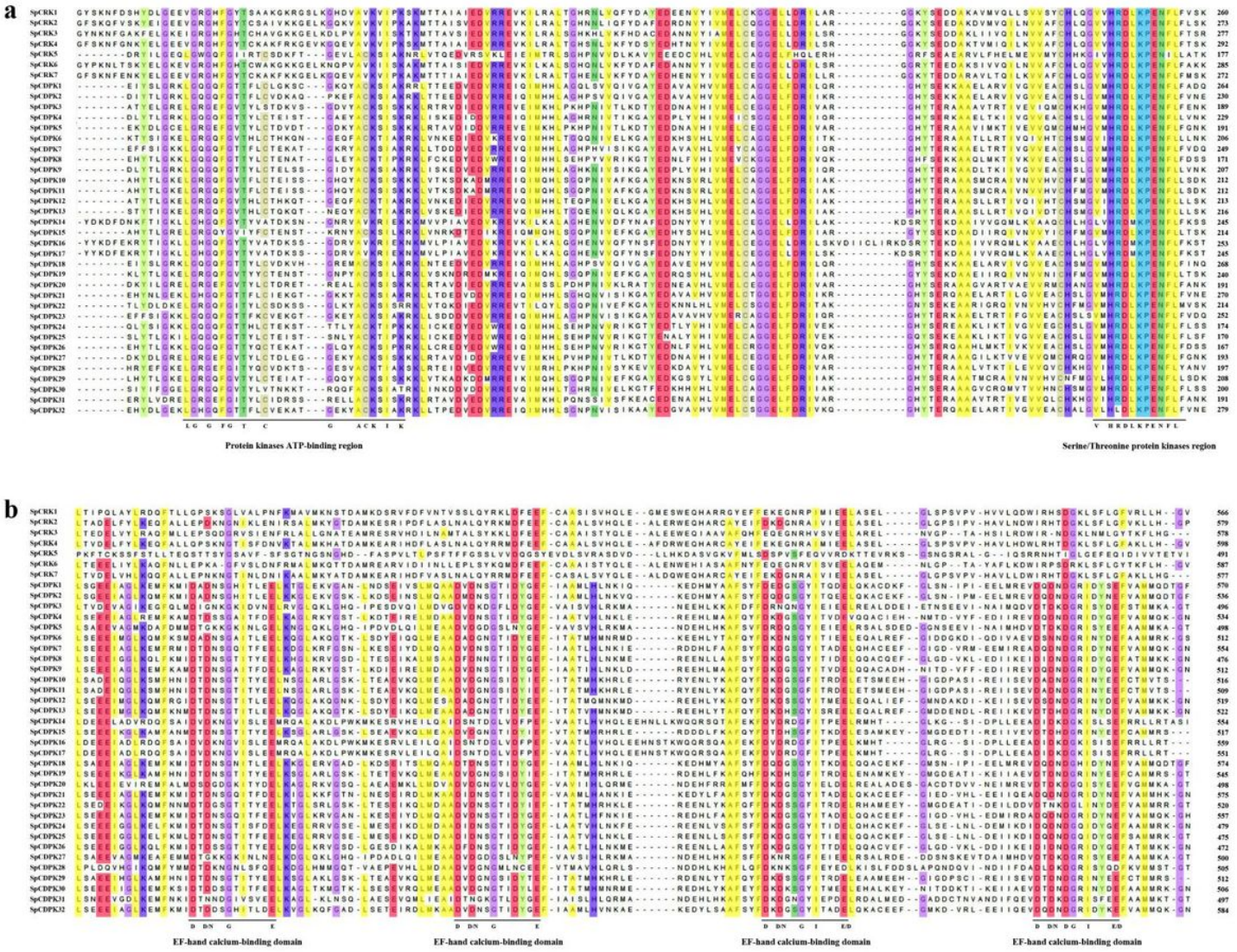

Figure 2

Comparison of conserved domains of SpCRK and SpCDPK gene families. Colour bars were used to mark no less than 70% of conserved sites.

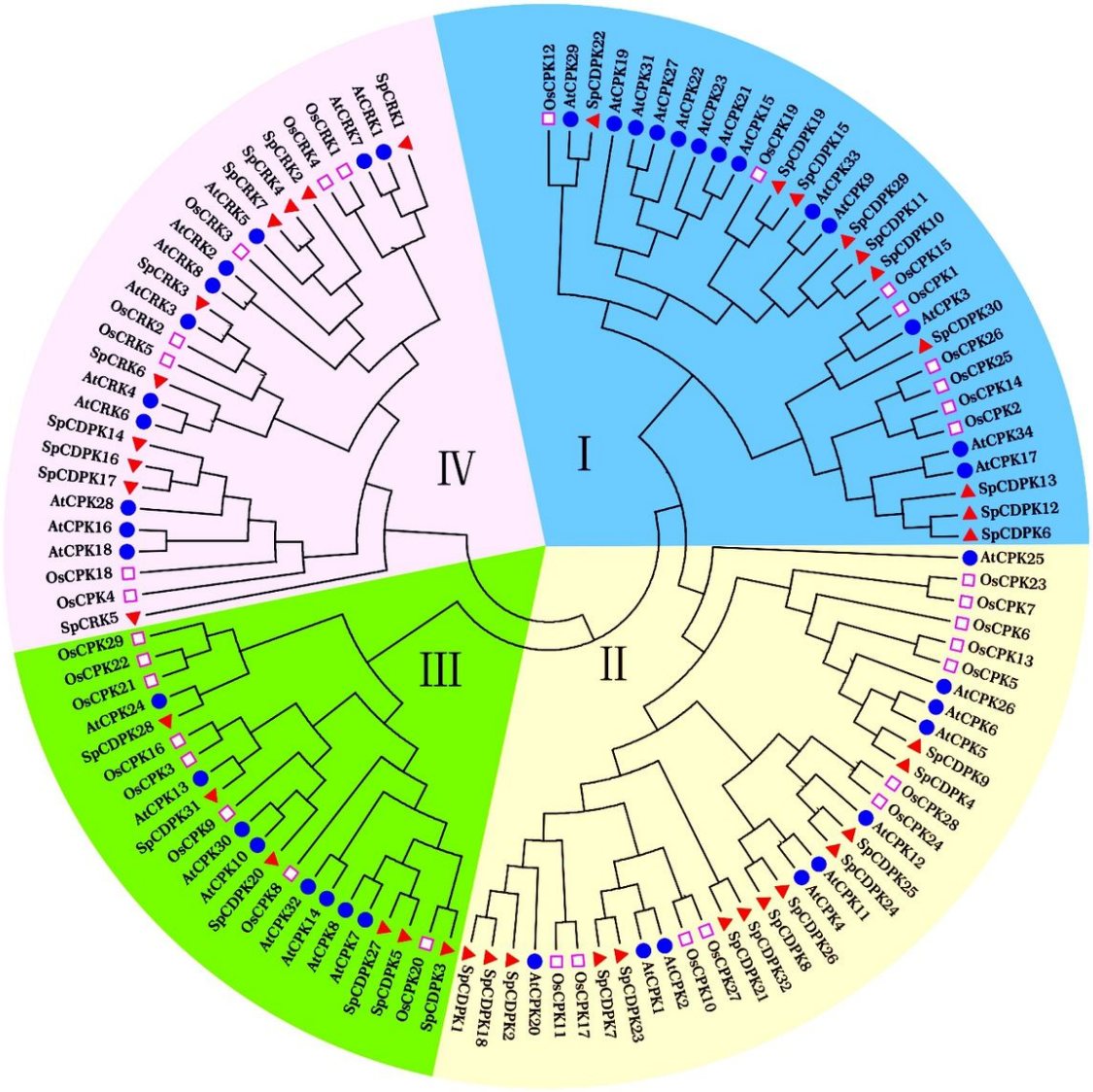

Figure 3

Phylogenetic relationship among CRK and CDPK family members from Solanum Pennellii, rice and Arabidopsis. To identify the plant species origin of each CRK and CDPK, a species acronym was included before the protein name: eg. SpCRK indicated CRK from Solanum Pennellii, AtCRK indicated CRK from Arabidopsis and OsCRK indicated CRK from rice. The red triangle, blue dots and pink border square before the protein names indicated CRK and CDPK from Solanum Pennellii, Arabidopsis and rice, respectively.

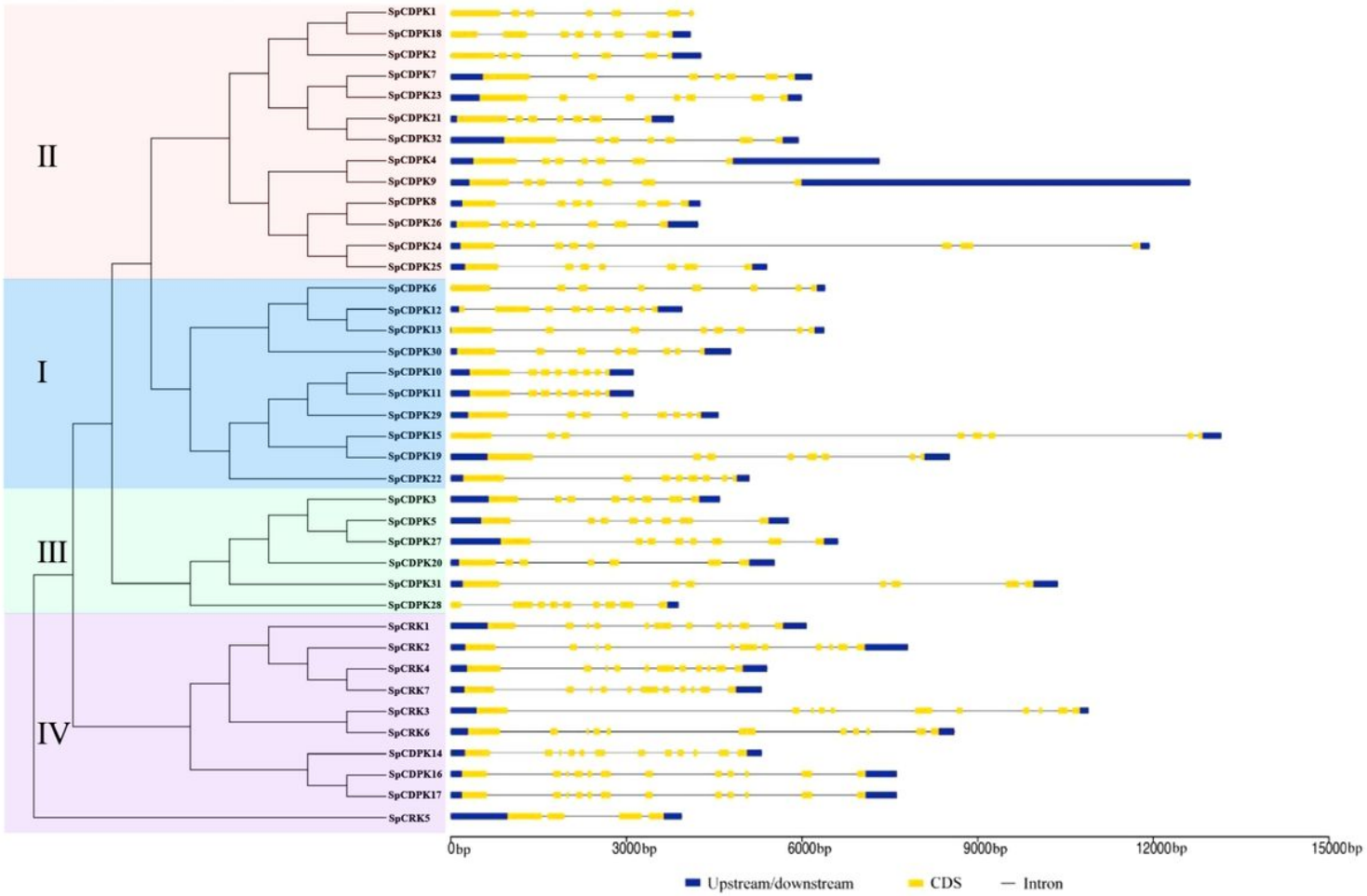

Figure 4

Phylogenetic tree and exon-intron structure of SpCRK and SpCDPK genes. The phylogenetic tree was constructed using the full-length protein sequences of 7 SpCRK and 32 SpCDPK. Introns and exons of the SpCRK and SpCDPK family genes were grouped according to the phylogenetic classification. Upstream/downstream, exons and introns were represented by blue boxes, yellow boxes, and the black lines respectively. The Exons and introns of SpCRK and SpCDPK genes was analyzed with Online website GSDS.

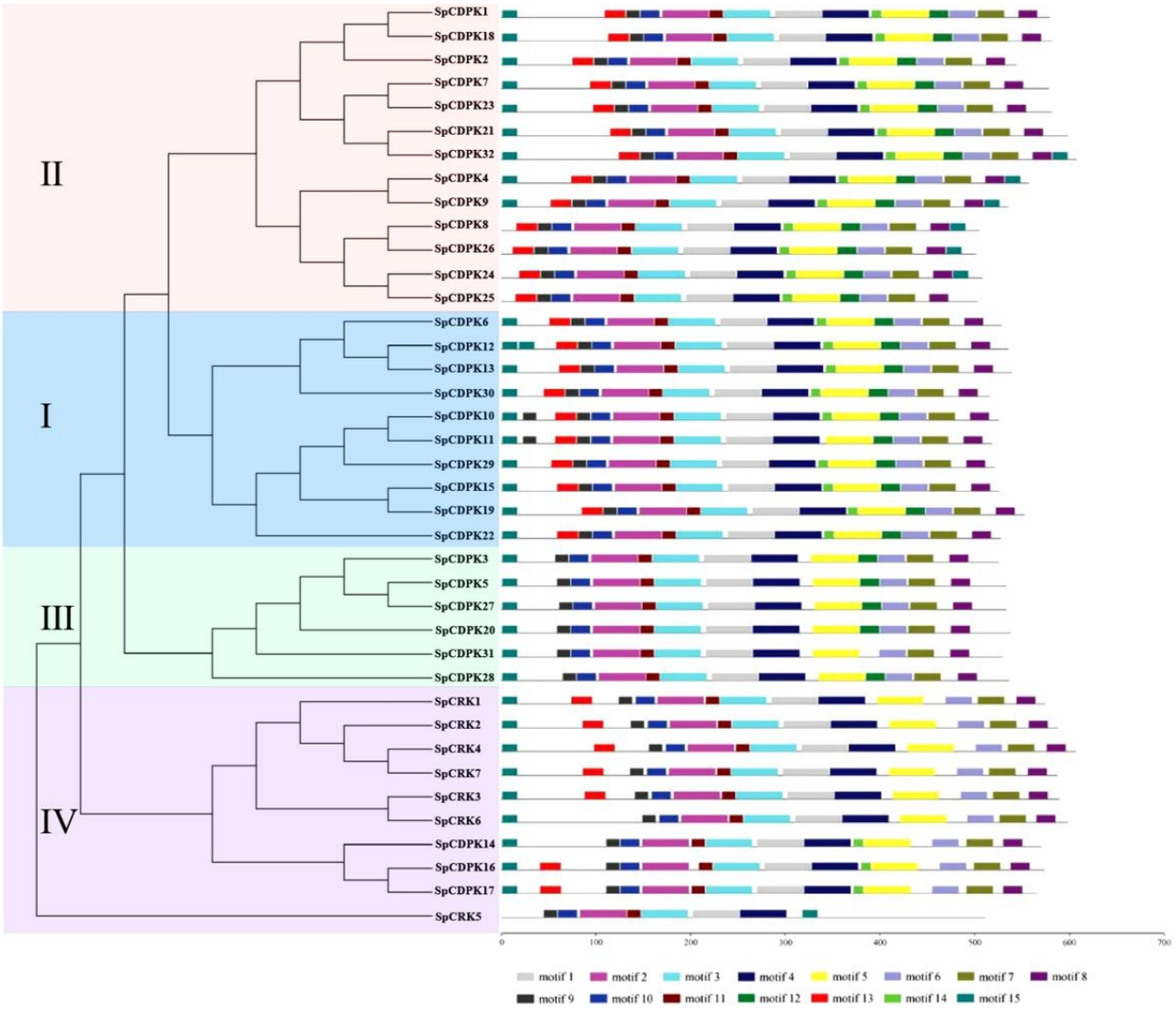

Figure 5

Phylogenetic tree and conserved motifs of SpCRK and SpCDPK proteins. The phylogenetic tree was constructed using the full-length protein sequences of 7 SpCRK and 32 SpCDPK. The conserved motifs of SpCRK and SpCDPK family proteins were grouped according to the phylogenetic classification. All motifs of SpCRK and SpCDPK proteins with complete amino acid sequences were identified by MEME database.

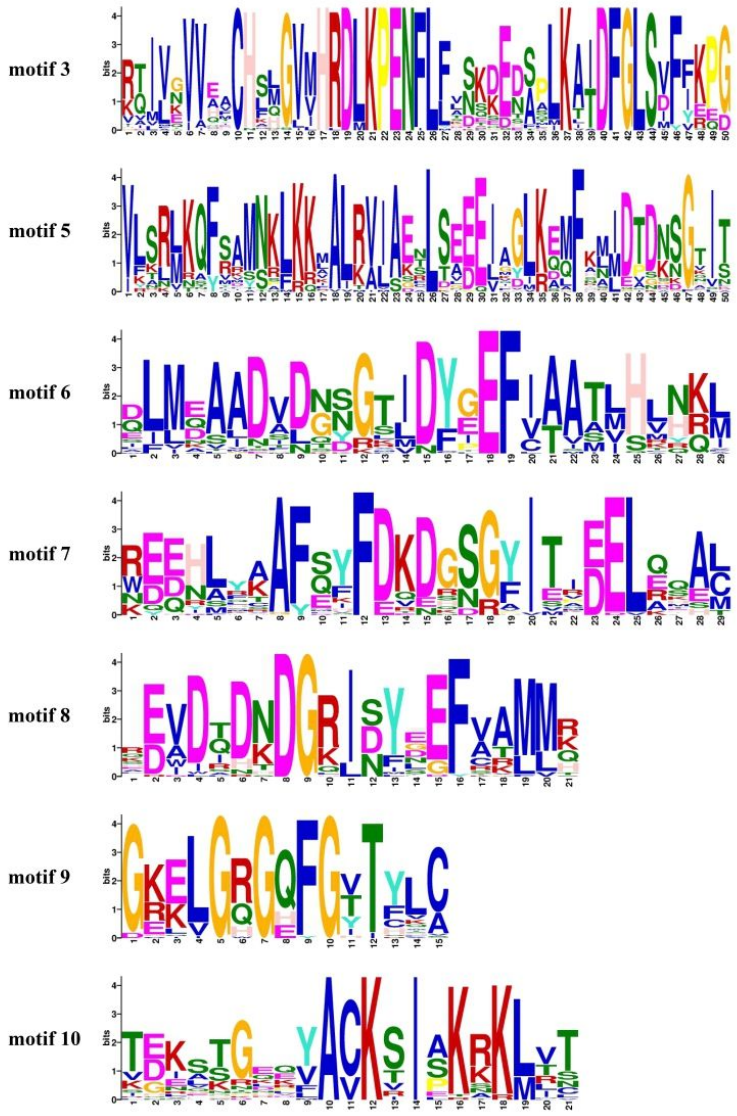

Figure 6

Motif LOGO. Motif 3 were annotated as Ser/Thr protein kinase region, and motifs 5, 6, 7 and 8 as EF-hand calcium-binding domain, and motifs 9 and 10 as protein kinases ATP-binding region.

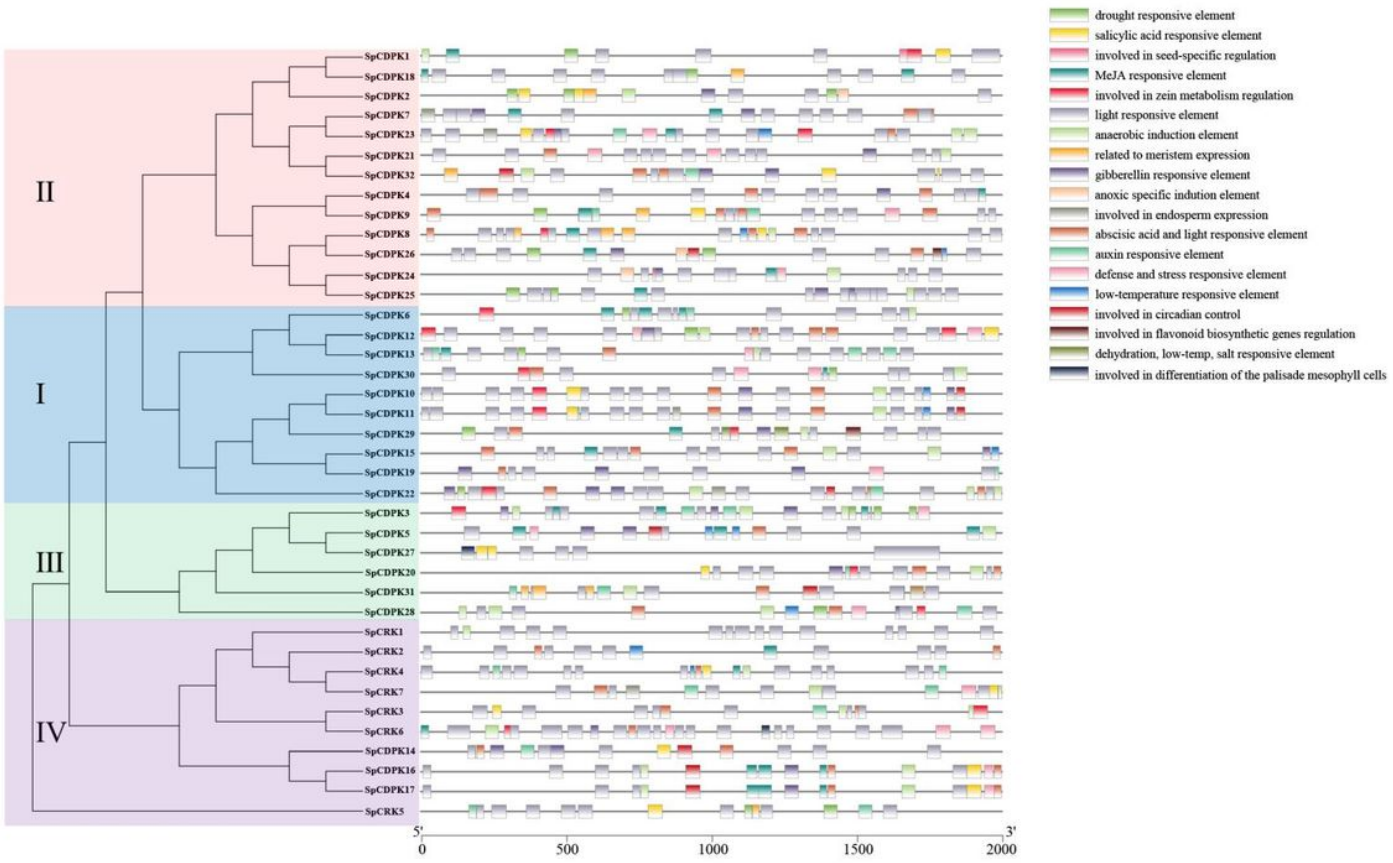

Figure 7

The cis-element in the promoter region of SpCRK and SpCDPK genes. The phylogenetic tree was constructed using the full-length protein sequences of 7 SpCRK and 32 SpCDPK. The cis-element in the promoter region of SpCRK and SpCDPK family genes were grouped according to the phylogenetic classification. All cis-elements in the promoter region of the upstream 2000bp sequence of SpCRK and SpCDPK genes were identified online by PlantCARE.