Chelonian Conservation And Biology



Vol. 18 No. 2 (2023) | https://www.acgpublishing.com/ | ISSN - 1071-8443 DOI: doi.org/10.18011/2023.11(2).1424.1434



Srikanth Burra<sup>1</sup>, K. Venkata Naga Rajeshwari<sup>1</sup>, S. Suma\*<sup>1</sup>

Department of Biochemistry, Chaitanya (Deemed to be University), Kishanpura, Hanumakonda, Telangana, India-506001 \*E-mail: sumareddy197420@gmail.com

Abstract: Plant growth and development, stress responses, senescence, and plant hormone signaling are just a few of the physiological processes for which calcium-dependent protein kinases (CDPKs) have been demonstrated to be crucial. We found 40 CDPK genes in our study by doing a bioinformatics analysis of the whole maize genome. Based on the phylogenetic study, 40 ZmCPKs are classified into four categories. we report a genome-wide analysis of the CDPK gene family in maize here. This genomic analysis of maize CDPK genes serves as a foundation for a functional investigation of this gene family in maize.

Keywords: CDPK, Gene Family, Maize, phylogenetic analysis.

# Introduction

Different mechanisms have evolved by plants to help them adapt to a range of harsh conditions, including biotic and abiotic pressures. Here, as the second messenger, calcium  $(Ca2+)$ engages in a variety of signal transduction pathways [2]

In response to brief changes in calcium concentration, calcium sensors or calcium-binding proteins can change the transcript of genes downstream, the phosphorylation of proteins, or the activity of enzymes (Sanders et al., 1999; Harmon et al., 2000). Plants have been shown to have three distinct kinds of calcium sensors thus far: calcineurin B-like protein (CBL), calciumdependent protein kinases (CDPKs), and calmodulin (CaM) and CaM-like protein (CaML). Only CDPKs, out of all these calcium sensors, have the unique ability to directly sense, respond, and translate Ca2+ signals into downstream protein phosphorylation without requiring a conformational change in the protein partner [19][11][20]. One of the well-known Ca2+-sensor protein kinases involved in environmental stress tolerance is the CDPK in plants. Four domains of the CDPK protein have been identified: the calmodulin-like domain, the autoregulatory/autoinhibitory domain, the catalytic domain of Ser/Thr kinase, and the N-terminal variable region [9][6][15]. The Ca2+ binding EF-hands are present in the calmodulin-like domain.



All the articles published by Chelonian Conservation and Biology are licensed under a Creative Commons Attribution-NonCommercial 4.0 International License Based on a work at https://www.acgpublishing.com/

CrossMark

According to Hamel et al. (2014) [8], the basal architecture of the CDPK family is extensively preserved in land plants, including bryophytes and pteridophytes. The four characterized regions of their conserved molecular structure are the autoinhibitory junction domain, calmodulin-like domain, N-terminal domain, and Ser/Thr kinase domain [12]. N-terminal domain length varies among species, and even within species among distinct CDPK variations [14]. This variation is linked to *in vivo* phosphorylation and protein localization. A key component of the CDPK function is the catalytic domain containing the ATP binding site, the Ser/Thr kinase domain.

Calcium concentrations inside cells have the potential to dramatically alter sub-cellular localization [21]. CDPK activity [10][2] and interactions with other proteins [13][25]. The CDPK-SnRK superfamily's autoinhibitory junction domain is thought to function as an auto-inhibitor, keeping CDPK inactive or allowing it to be stimulated by common cytosolic calcium signals. This allows CDPK activation to occur as a means of overcoming autoinhibition [25][20]. To sense calcium signals and bind to the free calcium ion, the C-terminal calmodulin-like domain of CDPKs contains up to four elongation factor (EF) hands, a 29 aa helix-loop-helix structure with 13 conserved residues.

CDPKs are a multigene family of plant-specific protein kinases. 34 CDPK genes are found in Arabidopsis (Arabidopsis thaliana; [5][12]), 29 in rice (Oryza sativa; [1]), 20 in wheat (Triticum saestivum; [15]), 40 in maize (Zea mays; [16]), 30 in poplar (Populus trichocarpa; [25]), 17 in grapevine (Vitis vinifera; [4]), respectively. The architecture of the CDPK family showed significant conservation from bryophytes to higher taxa, according to a long-term evolutionary perspective [8].

The genomes of rice, wheat, and Arabidopsis contain 31 genes, 34 CDPK genes, and 20 genes, respectively [22, 5,7]. Nicotiana tabacum, the tobacco plant, has numerous CDPK genes, as does Glycine max, the soybean plant, and Lycopersicon esculentum, the tomato plant [3,24]. But at the moment, not much is understood about the maize CDPK family. One of the most significant and ancient crops in the world is maize (Zea mays L.). It has only been possible to characterize seven ZmCPKs in maize thus far: ZmCK1, ZmCDPK1, ZmCDPK2, ZmCDPK7, ZmCDPK9, ZmCDPK10, and ZmCDPK11.

In this work, we found 40 CDPK genes by bioinformatics analysis of the whole maize genome. Based on their evolutionary connections, these 40 maize CDPKs have been categorized into 4 groups.

# Materials and Methods:

## CDPK Gene Identification by Database Search and Characterization in Maize

The protein sequences for thirty rice CDPKs and thirty Arabidopsis CDPKs were obtained from GenBank (http://www.ncbi.nlm.nih.gov/genbank/), the rice genome annotation database (http://rice.plantbiology.msu.edu/), and the Arabidopsis Information Resource (http://www.Arabidopsis.org/), respectively. The following link was used to retrieve sequences from the maize genome database: http://www.maizesequence.org/index.html. The Arabidopsis and rice CDPK protein sequences were utilized to scan the maize genome and NCBI database using BLASTP to identify the maize CDPK gene family. To get rid of duplication, a self-BLAST of the sequences was done. To validate the existence of the protein kinase domain, all potential candidates underwent manual verification using the InterProScan programme (http://www.ebi.ac.uk/Tools/pfa/iprscan/). Ultimately, all acquired protein sequences were subjected to additional analysis through the utilization of the Pfam (http://pfam.sanger.ac.uk/search) and SMART (http://smart.embl-heidelberg.de/) databases.

## Phylogenetic Analysis of CDPKs in Maize

Using the default settings of the Clustal X programme, calcium-dependent protein kinases from Arabidopsis, rice, and maize were aligned (Larkin *et al.*, 2007). The maximum-like hood approach was used to construct the phylogenetic tree under the MEGA 7.0 programme [23].

# Gene Duplication of CDPK genes in Maize

In maize B73, gene duplication events involving CDPK genes were examined. We used the following standards to define the gene duplication: 1) The aligned section had an identity >80%, 2) The alignment length spanned >80% of the larger gene, and 3) For closely connected genes, only one duplication event was recorded. Clustal X was used to align all of the pertinent genes found in the maize genomes, and MEGA v7.0 was used to compute the results.

#### Results and Discussion

# CDPK Gene Identification and Characterization in Maize

The entire sequencing of the maize genome has made it possible to identify every member of the CDPK gene family in that crop. To find 40 probable CDPK genes, including 7 known CDPKs that are referred to as ZmCPK1–ZmCPK40 following the suggested nomenclature for CDPK genes, BLAST searches of the maize genome were conducted using Arabidopsis and rice CDPK sequences as query sequences (Table 1). Based on the results of multiple sequence alignment and phylogenetic analysis, we found that alternative splice variants are closely linked to one another, therefore we decided to focus our investigation on only one variant. The total quantity of CDPK genes in maize was comparable to that of rice (389 Mb) and Arabidopsis (125 Mb), although the maize genome (2365 Mb) was substantially bigger. Furthermore, the increase of Group I, which accounted for 17 genes in maize, 11 in rice, and 10 in Arabidopsis, was the primary cause of the variation in the total number of CDPK genes (Figure 1).

The N-terminal variable domain, the protein kinase domain, the autoinhibitory domain, and the calmodulin-like domain were among the conserved CDPK domains. With a Gly residue at the second position, several CDPKs in Arabidopsis, rice, and wheat have potential N-myristoylation motifs for membrane attachment at the start of their extremely varied N-terminal domain. Nmyristoylation motifs were predicted for membrane attachment in 17 out of the 40 maize CDPKs. Fifteen CDPKs were found to possess a minimum of one Cys residue at positions 3, 4, or 5, which are putative locations for palmitoylation (Table 1). N-myristoylation and palmitoylation facilitate protein-membrane interactions in a variety of systems.

Plasma membranes have been linked to AtCPK2, AtCPK10, AtCPK3, TaCPK2, and TaCPK5, which have been predicted to have an N-myristoylation motif [24]. Additionally, it has been demonstrated through experimentation that OsCPK19 is myristoylated and palmitoylated before being directed towards the membrane fraction. These genes were also linked to the membranes, even though AtCPK5, AtCPK6, TaCPK3, and TaCPK15 lacked myristoylation patterns. Although ZmCK1 (ZmCPK3) was predicted to have an N-myristoylation motif, a ZmCK1:hGFP fusion protein was recently found to localize to the cytoplasm and nucleus .

The majority of CDPKs' calmodulin-like domains in wheat, rice, and Arabidopsis have four Ca2+ binding EF hands, which enable the protein to act as a  $Ca^{2+}$  sensor. The number and location of EF-hands may be significant for defining the  $Ca^{2+}$  regulation of CDPK activity, as shown by the sequential deletion of the EF hands. More recently, Franz et al. (2011) demonstrated that the N-terminal EF1 and EF2 motifs of AtCPK21 were more significant for  $Ca^{2+}$  regulated enzyme activity in comparison to the C-terminal EF3 and EF4 motifs. They also proposed that the EF1 and EF2 motifs could act as a switch for the protein kinase activity that mediates abiotic stress signaling. The majority of maize CDPKs had four EF hands; ZmCPK29, ZmCPK30, and ZmCPK31 each had three EF-hands (Table 1). Interestingly, ZmCPK7 only possessed one EFhand, a trait shared by its homologs AtCPK25 and OsCPK7. Examining the variations in ZmCPK7's and other ZmCPKs' biological roles would be interesting.

## Phylogenetic Analysis of CDPK Gene Family

Using alignments of the complete CDPK amino acid sequences, an unrooted tree was built to investigate the evolutionary relationships amongst various CDPK members. The phylogenetic analysis revealed that 40 ZmCPKs can be categorized into four groups (Figure 1). According to reports, CRKs, PPCKs, and PEPRKs have a close relationship with CDPKs. ZmCPKs, AtCRKs, AtPPCKs, and AtPEPRKs were grouped into 7 different groups, as seen in Figure 1, indicating that all 40 ZmCPKs genuinely belonged to the CDPK family. Eleven CDPKs from rice, ten from Arabidopsis, and seventeen from maize made up Group I. OsCPK13 (OsCDPK7), which has been linked to reactions to cold, drought, and salt stress, shared a lot of similarities with ZmCPK1, ZmCPK2, ZmCPK3, and ZmCPK4.

Similarities suggest that ZmCPK1, ZmCPK2, ZmCPK3, and ZmCPK4 may have an abiotic stress-related role. Eight rice CDPKs, thirteen Arabidopsis CDPKs, and eleven maize CDPKs were in Group II. The positive regulator of salt tolerance and the negative regulator of blast resistance, OsCPK12, and ZmCPK25, had 76% of the same similarities. There were eight CDPKs from rice, nine from maize, and eight from Arabidopsis in Group III. ZmCPK32 and OsCPK21's amino acid sequences were around 75% identical. It's been demonstrated that rice can tolerate salt thanks to OsCPK21. Three maize CDPKs, three rice CDPKs, and three Arabidopsis CDPKs were present in Group IV. On the level of amino acids, ZmCPK38 and AtCPK28 were 75% identical. Based on phylogenetic analysis, it was determined that there were 14 closely related orthologs between maize and rice (ZmCPK5 and OsCPK5, ZmCPK6 and OsCPK6, ZmCPK7 and OsCPK7, ZmCPK10 and OsCPK27, ZmCPK11, and OsCPK10, ZmCPK17 and OsCPK28, ZmCPK18 and OsCPK2, ZmCPK19 and OsCPK14, ZmCPK22, and OsCPK15, ZmCPK25 and OsCPK12, ZmCPK31 and OsCPK3, ZmCPK32 and OsCPK21, ZmCPK38 and OsCPK18) (Figure 1), indicating the existence of an ancestral set of CDPK genes before the divergence between maize and rice.

## Conclusion

One of the biggest families of protein kinases in plants, the CDPK gene family, was examined genome-wide in maize in this work. A thirty-three-strong list of newly identified CDPK genes from maize was found by in silico analysis of the maize genome database. According to our analysis, ZmCPK growth might have been facilitated by genome duplications. Four groups of these genes have been formed thanks to the phylogenetic analysis of CDPKs from Arabidopsis, rice, and maize. Due to the presence of conserved protein motifs, members of each group may have recently evolved together. All things considered, we have identified the CDPK family, and the data we have shared here offers a strong basis for additional functional research into the maize CDPK gene family.



Figure 1: Phylogenetic tree of CDPKs from Maize, rice, and Arabidopsis. Neibour-joining tree was created using MEGA 7.0 Program with 1000 bootstraps using full-length sequences of 40

maize, 29 rice, and 34 Arabidopsis CDPK proteins. Four groups were labeled as I, II, III, and IV.

Note: Red boxes: maize-rice orthologs, blue boxes: duplicated genes.









Note: a, b, c are GenBank accession numbers

#### References:

- 1. Asano T, Tanaka N, Yang G, Hayashi N, Komatsu S: 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 (2005), 46:356–366.
- 2. Boudsocq, M., Droillard, M. J., Regad, L., and Lauriere, C. (2012). Characterization of Arabidopsis calcium-dependent protein kinases: activated or not by calcium? Biochem. J. 447, 291–299. doi: 10.1042/Bj20112072.
- 3. Chang, W. J., Su, H. S., Li, W. J., and Zhang, Z. L. (2009). Expression profiling of a novel calcium-dependent protein kinase gene, LeCPK2, from tomato (Solanum lycopersicum) under heat and pathogen-related hormones. Biosci. Biotechnol. Biochem. 73, 2427–2431. doi: 10.1271/Bbb.90385.
- 4. Chen, F., Fasoli, M., Tornielli, G. B., Dal Santo, S., Pezzotti, M., Zhang, L., et al. (2013).
- 5. Cheng, S. H., Willmann, M. R., Chen, H. C., and Sheen, J. (2002). Calcium signaling through protein kinases. The Arabidopsis calcium-dependent protein kinase gene family. Plant Physiol. 129, 469–485. doi: 10.1104/Pp.005645
- 6. Christodoulou J, Malmendal A, Harper JF, Chazin WJ: Evidence for differing roles for each lobe of the calmodulin-like domain in a calcium-dependent protein kinase. J Biol Chem (2004), 279:29092–29100.
- 7. Geng S, Zhao Y, Tang L, Zhang R, Sun M, Guo H, Kong X, Li A, Mao L: Molecular evolution of two duplicated CDPK genes CPK7 and CPK12 in grass species: a case study in wheat (Triticum aestivum L). Gene 2011, 475:94–103.
- 8. Hamel, L. P., Sheen, J., and Sequin, A. (2014). Ancient signals: comparative genomics of green plant CDPKs. Trends Plant Sci. 19, 79–89. doi: 10.1016/j.tplants.2013.10.009
- 9. Harper JF, Sussman MR, Schaller GE, Putnam-Evans C, Charbonneau H, Harmon AC: A calcium-dependent protein kinase with a regulatory domain similar to calmodulin. Science 1991, 252:951–954.

10. Harper, J. E., Breton, G., and Harmon, A. (2004). Decoding  $Ca^{2+}$  signals through plant protein kinases. Ann. Rev. Plant Biol. 55, 263–288. doi: 10.1016/j.tplants.2013.10.009

- 11. Harper, J. F., and Harmon, A. (2005). Plants, symbiosis and parasites: a calcium signalling connection. Nat. Rev. Mol. Cell Biol. 6, 555–566. doi: 10.1146/annurev.arplant.55.031903.141627.
- 12. Hrabak, E. M., Chan, C. W. M., Gribskov, M., Harper, J. F., Choi, J. H., Halford, N., et al. (2003). The Arabidopsis CDPK-SnRK superfamily of protein kinases. Plant Physiol. 132, 666–680. doi: 10.1104/pp.102.011999.
- 13. Ishida, S., Yuasa, T., Nakata, M., and Takahashi, Y. (2008). A tobacco calcium-dependent protein kinase, CDPK1, regulates the transcription factor
- 14. Ito, T., Nakata, M., Fukazawa, J., Ishida, S., and Takahashi, Y. (2010). Alteration of substrate specificity: the variable N-terminal domain of tobacco  $Ca^{2+}$ -dependent protein kinase is important for substrate recognition. Plant Cell 22, 1592–1604. doi: 10.1105/tpc.109.073577.
- 15. Klimecka M, Muszynska G: Structure and functions of plant calcium-dependent protein kinases. Acta Biochim Pol 2007, 54:219–233.
- 16. Kong, X. P., Lv, W., Jiang, S. S., Zhang, D., Cai, G. H., Pan, J. W., et al. (2013). Genomewide identification and expression analysis of calcium-dependent protein kinase in maize. BMC Genomics 14:433. doi: 10.1186/1471-2164-14-433.
- 17. Li A, Wang X, Leseberg CH, Jia J, Mao L: Biotic and abiotic stress responses through calciumdependent protein kinase (CDPK) signaling in wheat (Triticum aestivum L). Plant Signal Behav 2008, 3:654–656.
- 18. Li, A. L., Zhu, Y. F., Tan, X. M., Wang, X., Wei, B., Guo, H. Z., et al. (2008). Evolutionary and functional study of the CDPK gene family in wheat (Triticum aestivum L.). Plant Mol. Biol. 66, 429–443. doi: 10.1007/s11103-007-9281-5
- 19. Ludwig, A. A., Saitoh, H., Felix, G., Freymark, G., Miersch, O., Wasternack, C., et al. (2005). Ethylene-mediated cross-talk between calcium-dependent protein kinase and MAPK signaling controls stress responses in plants. Proc. Natl. Acad. Sci. U.S.A. 102, 10736–10741. doi:10.1073/pnas.0502954102
- 20. Poovaiah, B. W., Du, L. Q., Wang, H. Z., and Yang, T. B. (2013). Recent advances in calcium/calmodulin-mediated signaling with an emphasis on plant-microbe interactions. Plant Physiol. 163, 531–542. doi: 10.1104/pp.113.220780
- 21. Raichaudhuri, A., Bhattacharyya, R., Chaudhuri, S., Chakrabarti, P., and Dasgupta, M. (2006). Domain analysis of a groundnut calcium-dependent protein kinase - Nuclear localization sequence in the junction domain is coupled with nonconsensus calcium-binding domains. J. Biol. Chem. 281, 10399–10409. doi: 10.1074/jbc.M511001200
- 22. Ray S, Agarwal P, Arora R, Kapoor S, Tyagi AK: Expression analysis of calcium-dependent protein kinase gene family during reproductive development and abiotic stress conditions in rice (Oryza sativa L ssp Indica). Mol Genet Genomics (2007), 278:493–505.REPRESSION OF SHOOT GROWTH in response to gibberellins. Plant Cell 20, 3273–3288. doi: 10.1105/tpc.107.057489
- 23. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739. doi: 10.1093/molbev/msr121
- 24. Yang, T. B., and Poovaiah, B. W. (2003). Calcium/calmodulin mediated signal network in plants. Trends Plant Sci. 8, 505–512. doi: 10.1016/j.tplants.2003.09.004
- 25. Zou, J. J., Wei, F. J., Wang, C., Wu, J. J., Ratnasekera, D., Liu, W. X., et al. (2010). Arabidopsis calcium-dependent protein kinase CPK10 functions in abscisic acid- and  $Ca^{2+}$ -mediated stomatal regulation in response to drought stress. Plant Physiol. 154, 1232–1243. doi: 10.1104/pp.110.157545