

Chapter 2

Genomic Organization

Abstract After the completion of several plant genomes, *CBL* and *CIPK* genes have been identified in various plant species beside *Arabidopsis* and rice in angiosperms. Moreover, *CBL* and *CIPK* genes have also been identified in algae, mosses, pteridophytes, and gymnosperms. Interestingly, CBL- and CIPK-type proteins have also been identified outside plant kingdom in protozoa. But, very little is known about the functional role of the CBL and CIPK in these species.

Keywords Genome • Organization • Evolution • CBLs • CIPKs • Gene • Phylogenetic

2.1 Introduction

As organisms become more complex, more numbers of attributes are added to its function. Specificity and cross talks are the most important part of any signaling network in complex genomes; and therefore, it is an interesting question to decipher how signaling molecules and networks are affected during the course of evolution in organisms. In order to answer this question, it is important to understand the genomic organization of a signaling network. CBL and CIPK genes have been identified in algae, mosses, pteridophytes, gymnosperms, and various species of angiosperms (Fig. 2.1) [1, 4, 5, 8, 11, 12]. CIPK constitutes a large gene family in plant genomes compared to CBL. The number of *CBL* and *CIPK* genes increases in the respective gene families from lower plants to higher plants [8]. These findings also suggest that the evolution of the plant lineage coincided with the evolution of complexity of the CBL and CIPK gene families. Duplication of genes played very significant role in the expansion of this gene family. In this chapter, the genomic organization of both *CBL* and *CIPK* gene(s) will be addressed.

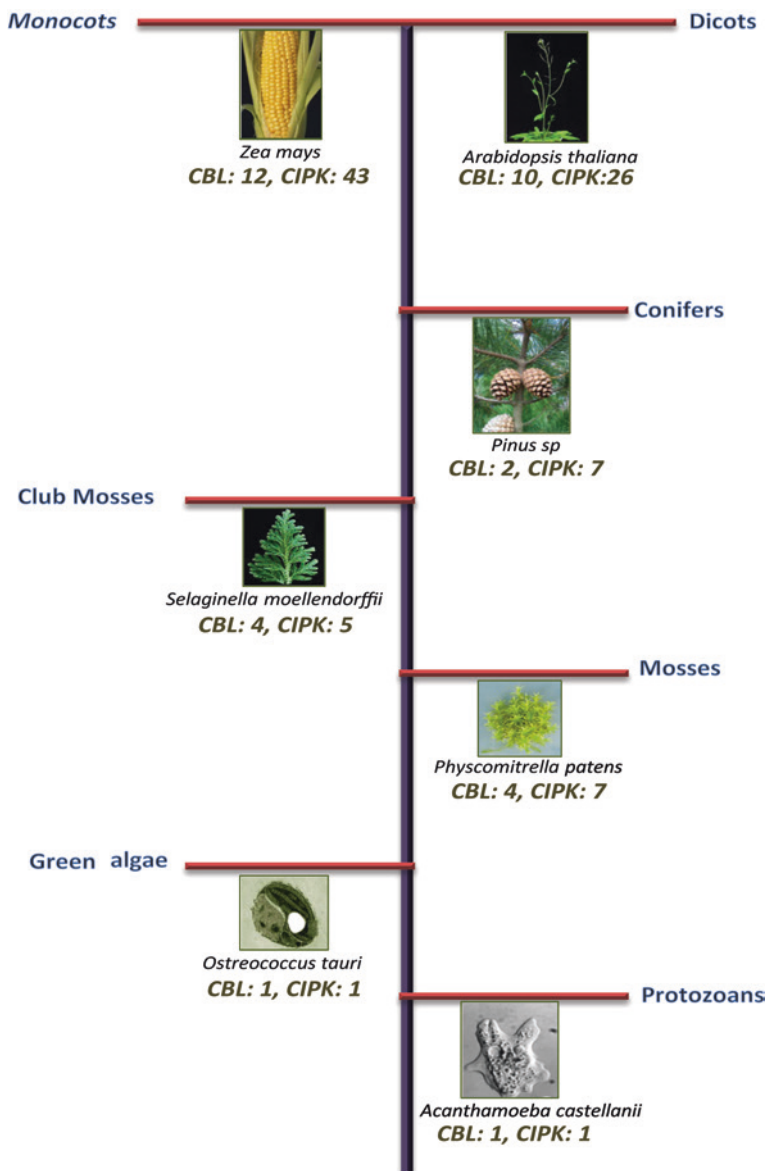


Fig. 2.1 Evolution and complexity of CBLs and CIPKs protein families in different plant lineages. Number of CBL and CIPK genes increases in the respective gene families from lower plants to higher plants suggesting the complexity of the CBL–CIPK system concurrently evolved with the increasing morphological and developmental sophistication of plants

2.2 CBL and CIPK Complements

2.2.1 CBL- and CIPK-Type Proteins in Protozoan

Though, a quest for CBL- and CIPK-related protein in different organism lead to the identification of the calcium signaling module in protozoan eukaryotes such as *Naegleria gruberi* and *Trichomonas vaginalis* [2], but phylogenetic analysis related them more to plant CBL proteins in comparison with animal CNB- and NCS-related proteins [2, 8]. EF-hand 1 was a common structure between protozoans CBL-like and the plants CBLs, similarly, protozoan CIPKs have the characteristic NAF domain-like plant CIPKs. With availability of more genome sequences, CBL and CIPK proteins are also identified in other protozoans like *Acanthamoeba castellanii* and *Emiliania huxleyi* CCMP1516. The presence of both protein families in these non-plant species provides clues for the evolution of CBL–CIPK signaling components.

2.2.2 CBL and CIPK in Algae

A single CBL and CIPK sequences have been identified in single-celled organisms such as *Chlorella* and *Ostreococcus* [2, 8], suggesting early evolution of the CBL–CIPK system at the base of plant lineage. Beside this, CBL–CIPK homologs were discovered from other charophyte green algal species including *Coleochaete orbicularis*, *Lebsormidium flaccidum*, *Chaetosphaeridium globosum*, and *Penium margaritaceum*. Charophyte CBL and CIPK homologs displayed expected domain architecture and motifs like their angiosperm counterpart, but these proteins were found to be absent in the algae *Volvox carterii* and *Chlamydomonas reinhardtii*. The exact reason of why CBLs and CIPKs are absent in these algae is not clear; however, these organism might have acquired other Ca^{2+} signaling system for transducing the similar stimuli [2, 8].

2.2.3 CBL and CIPK in Higher Plants

Physcomitrella patens encode four CBL and seven CIPK proteins [8]. The fern *Selaginella moellendorffii* possesses a complement of four CBLs and five CIPKs [8], whereas conifers like *Pinus* sp. consists of two CBL proteins and seven CIPKs [5]. CBL and CIPK proteins were originally identified in the dicot plant *Arabidopsis thaliana* [6, 7]. After the availability of genome sequence of *Arabidopsis*, a comprehensive bioinformatic analyses identified a complement of 10 CBLs and 26 CIPKs [5, 8]. Subsequently, availability of genome sequence in other plants also yield to identify multiple members of CBL–CIPK gene family i.e, 10 genes encoding CBL proteins and 27 genes encoding CIPKs in genome of

Populus trichocarpa (poplar) [8, 11, 12], 8 distinct *CBLs*, and 21 *CIPKs* in *Vitis vinifera* [8]. In monocots, a total of 10 *CBLs* and 30 *CIPKs* were identified in *Oryza sativa* [1, 5]. *Zea mays* has largest *CBL* and *CIPK* gene family members with 12 *CBLs* and 43 *CIPKs* in its genome [4]. However, only 6 *CBLs* and 32 *CIPK*-type kinases were identified in *Sorghum bicolor* [8].

2.3 Genomic Architecture

In the case of multi-gene family, chromosomal distributions of genes on different chromosome provide significant information about the evolution of any gene and gene family in the genome. In the case of genomic distribution of Arabidopsis *CBLs*, most of the genes are present on chromosomes IV and V, whereas in rice, *CBL* genes were located on 6 different chromosomes. As mentioned earlier that the number of *CBLs* and *CIPKs* in the genome increases from lower organism such as protozoa and green algae to higher plants, especially the number of *CIPKs* genes, so here the role of segmental and tandem duplications might have played a significant role in the expansion of *CIPK* gene family. In Arabidopsis, there are eight segmental duplication events and two *CIPK* gene pairs arose by tandem duplications [5]. These duplication events are responsible for generation of many homologous gene pairs in the *CIPK* gene family and amplification of this gene family. Similarly in rice also most of genes arose because of segmental and tandem duplications. According to evolutionary study done by Ye et al. [10] in angiosperm (*O. sativa*, *Z. mays*, *A. thaliana*, and *P. trichocarpa*), it was found that gene duplications has played an important role in the expansion of *CIPKs*.

2.4 Gene Structure

CBLs generally contain multiples introns. In Arabidopsis, rice, *S. bicolor* most of *CBL* genes harbor seven introns in their coding region. *CIPKs* are divided into two groups based on protein sequence similarity and number of introns, based on later criteria there are two types of *CIPK* gene groups: intron-rich and intron-poor groups. This kind of grouping has been observed in Arabidopsis, rice, and *Z. mays*. In *Z. mays* out of the total 43 *ZmCIPKs*, 12 *ZmCIPKs* belong to intron-rich group and rest are intron free [4]. In Arabidopsis, out of 25 *CIPK*, eight harbor multiple intron sequences. Rice and poplar also show similar kind of trend. Whereas, in the case of moss *P. patens*, all *CIPKs* contain multiple introns. Furthermore, all the *CIPKs* from the non-angiosperm species were found in intron-rich clade. Analysis of gene duplication showed that the expansion of *CIPKs* (intron-rich and intron-poor) is partly contributed by segmental duplications; however, tandem duplicates were found only in intron-poor *CIPKs* [10].

2.5 Phylogenetic Relatedness and Evolution

CBL number among species varied because of species-specific duplication or deletion genetic modifications [8]. The algal, fern, and moss CBL protein families harbor an N-terminal MGCXXS/T motif that have been shown to allow lipid modification by myristoylation and S-acylation of *A. thaliana* CBL1 [3]. During the evolution, the structural diversification leads to many changes in this membrane-anchored signaling module, which may have resulted in alternative subcellular localizations. Phylogenetic analysis [8] showed that the Arabidopsis CBL10 protein forms a distinct cluster with duplicated CBL protein pairs from poplar, grape, and rice. In general, the comparative analysis of CBLs from all these species further supports the classification of these proteins according to their N-terminal domains.

Ye et al. [10] generated a phylogenetic tree of 146 CIPKs from eight different plant genomes including *Chlorella*, *Ostreococcus tauri*, *P. patens*, *S. moellendorffii*, *O. sativa*, *Z. mays*, *A. thaliana*, and *P. trichocarpa* grouped into two clades, designated as intron-rich clade and intron-less clade. All the CIPKs proteins in algae, moss, and spike moss were grouped into intron-rich clade. According to phylogenetic analysis of various CIPKs of *P. patens*, *S. moellendorffii*, *A. thaliana*, and *S. bicolor*, all CIPKs from *P. patens* and *S. moellendorffii* were found to array together with CIPK23 and CIPK24 from *A. thaliana*, shown to represent critical regulators of plant K⁺ and Na⁺ homeostasis, respectively [9, 13]. This situation may reflect a functional conservation of the CBL–CIPK system in regulating the transport and distribution of these ions in mosses and ferns [8].

References

1. Albrecht V, Ritz O, Linder S, Harter K, Kudla J (2001) The NAF domain defines a novel protein-protein interaction module conserved in Ca²⁺-regulated kinases. *EMBO J* 20:1051–1063
2. Batistic O, Kudla J (2009) Plant calcineurin B-like proteins and their interacting protein kinases. *Biochim Biophys Acta* 1793:985–992
3. Batistic O, Sorek N, Schultke S, Yalovsky S, Kudla J (2008) Dual fatty acyl modification determines the localization and plasma membrane targeting of CBL/CIPK Ca²⁺ signaling complexes in Arabidopsis. *Plant Cell* 20:1346–1362
4. Chen X, Gu Z, Xin D, Hao L, Liu C, Huang J, Ma B, Zhang H (2011) Identification and characterization of putative CIPK genes in maize. *J Genet Genomics* 38:77–87
5. Kolukisaoglu U, Weinl S, Blazevic D, Batistic O, Kudla J (2004) Calcium sensors and their interacting protein kinases: genomics of the Arabidopsis and rice CBL–CIPK signaling networks. *Plant Physiol* 134:43–58
6. Kudla J, Xu Q, Harter K, Gruissem W, Luan S (1999) Genes for calcineurin B-like proteins in Arabidopsis are differentially regulated by stress signals. *Proc Natl Acad Sci USA* 96:4718–4723
7. Shi J, Kim KN, Ritz O, Albrecht V, Gupta R, Harter K, Luan S, Kudla J (1999) Novel protein kinases associated with calcineurin B-like calcium sensors in Arabidopsis. *Plant Cell* 11:2393–2405

8. Weinl S, Kudla J (2009) The CBL–CIPK Ca^{2+} -decoding signaling network: function and perspectives. *New Phytol* 184:517–528
9. Xu J, Li HD, Chen LQ, Wang Y, Liu LL, He L, Wu WH (2006) A protein kinase, interacting with two calcineurin B-like proteins, regulates K^+ transporter AKT1 in Arabidopsis. *Cell* 125:1347–1360
10. Ye CY, Xia X, Yin W (2013) Evolutionary analysis of CBL-interacting protein kinase gene family in plants. *Plant Growth Regul* 71:49–56
11. Yu Y, Xia X, Yin W, Zhang H (2007) Comparative genomic analysis of CIPK gene family in Arabidopsis and Populus. *Plant Growth Regul* 52:101–110
12. Zhang H, Yin W, Xia X (2008) Calcineurin B-Like family in Populus: comparative genome analysis and expression pattern under cold, drought and salt stress treatment. *Plant Growth Regulation* 56:129–140
13. Zhu JK (2003) Regulation of ion homeostasis under salt stress. *Curr Opin Plant Biol* 6:441–445

Global Comparative Analysis of CBL-CIPK Gene Families
in Plants

Pandey, G.; Kanwar, P.; Pandey, A.

2014, IX, 91 p. 19 illus. in color., Softcover

ISBN: 978-3-319-09077-1