Evolutionary Analysis of Calcium-Dependent Protein Kinase in Five Asteraceae Species

Liping Zhu 1,*, Bowen Zheng 1,†, Wangyang Song 1, Hongbin Li 1,2,* and Xiang Jin 1,2,*

1 Key Laboratory of Xinjiang Phytopharmacy Resource and Utilization of Ministry of Education, College of Life Sciences, Shihezi University, Shihezi 832003, China; zhuliping903@163.com (L.Z.); zhengbw0609@163.com (B.Z.); swywinner@163.com (W.S.)
2 Ministry of Education Key Laboratory for Ecology of Tropical Islands, College of Life Sciences, Hainan Normal University, Haikou 571158, China

* Correspondence: lihb@shzu.edu.cn (H.L.); jinx@hainnu.edu.cn (X.J.)
† These authors contribute equally to this work.

Received: 5 November 2019; Accepted: 21 December 2019; Published: 24 December 2019

Abstract: Calcium-dependent protein kinase (CPK) is crucial in Ca^{2+} signal transduction, and is a large gene family in plants. In our previous work, we reported Hevea brasiliensis CPKs were important for natural rubber biosynthesis. However, this CPK gene family in other rubber producing plants has not been investigated. Here, we report the CPKs in five representative Asteraceae species, including three rubber-producing and two non-rubber species. A total of 34, 40, 34 and 30 CPKs were identified from Taraxacum kokssaghzyaz, Lactuca sativa, Helianthus annuus, Chrysanthemum nankingense and Cynara cardunculus, respectively. All CPKs were classified into four individual groups (group I to IV). In addition, 10 TkCPKs, 11 LsCPKs, 20 HaCPKs, 13 CnCPKs and 7CcCPKs duplicated paralogs were identified. Further evolutionary analysis showed that, compared to other subfamilies, the group III had been expanded in the Asteraceae species, especially in the rubber-producing species. Meanwhile, the CPKs in group III from Asteraceae species tend to expand with low calcium binding capacity. This study provides a systematical evolutionary investigation of the CPKs in five representative Asteraceae species, suggesting that the subfamily specific expansion of CPKs might be related to natural rubber producing.

Keywords: calcium-dependent protein kinases; natural rubber producing; evolutionary analysis; Asteraceae; gene family expansion

1. Introduction

Calcium (Ca^{2+}) participates in miscellaneous signal transduction pathways as the second messenger such as stress, immune and signaling. To date, there are three major classes of Ca^{2+}-binding proteins that have been characterized in higher plants, including calcium-dependent protein kinases (CPK), calmodulins (CaMs) and CaM-like proteins (CaMLs), and calcineurin B-like proteins (CBLs) [1,2]. The CPK constitutes one of the largest protein kinase families that sense the calcium signal in plants [3]. The CPKs are monomeric proteins with structures that contain four conserved domains: the N-terminal variable domain, serine/threonine kinase domain, auto-inhibitory junction domain and the calmodulin-like domain [4,5]. The N-terminal domain is highly variable and contains myristoylation or palmitoylation sites for subcellular targeting [6]. The protein kinase domain is the catalytic domain with an adenosine triphosphate (ATP) binding site, which is often followed by the auto-inhibitory domain that serves as an auto-inhibitor to switch CPKs between inactive and/or active forms depending on the level of calcium concentration [7]. Moreover, the calmodulin-like domain often contains four EF-hands for Ca^{2+} binding [8,9].
2. Results

2.1. Identification of CPKs in Five Asteraceae Species

Genome-wide analyses of the Asteraceae species allowed us to identify 34, 34, 40, 34 and 30 CPK members in Tks, L. sativa, H. annuus, C. nankingense and C. cardunculus. Previously reported 34 At CPKs, 30 OsCPKs, 39 GmCPKs, 29 StCPKs, 26 StCPKs and 30 HbCPKs were obtained from corresponding databases [5,28–32].
Among the eleven species, the numbers of exons in CPK were not significantly different from each other in group I–III, but exon numbers in group IV were significantly higher than that in groups I–III (Table 1). Coding sequence lengths and encoded polypeptides of TkCPKs, CnCPKs, HaCPKs, LsCPKs and CcCPKs range of 1452–1779 bp/483–592 aa, 1230–1815 bp/409–604 aa, 1470–1896 bp/490–632 aa, 1464–1800 bp/488–600 aa and 1470–1824 bp/489–607 aa, respectively. In Asteraceae species, 22 TkCPKs, 11 CnCPKs, 25 HaCPKs, 20 LsCPKs and 18 CcCPKs contain predicted palmitoylation sites and 28 TkCPKs, 22 CnCPKs, 35 HaCPKs, 28 LsCPKs and 25 CcCPKs have putative myristoylation sites (Table 1). These CPKs may function in many physiological processes by membrane association in plants, because myristoylation or palmitoylation sites within the N-terminal variable domain can facilitate membrane association.

Table 1. Detailed information of calcium-dependent protein kinases (CPKs) in this study.

| CPKs  | Species                  | Exon Number | CDS    | aa     | Pal b | Myr b |
|-------|--------------------------|-------------|--------|--------|-------|-------|
| OsCPKs| Oryza sativa             | 3–8         | 1539–1839 | 513–613 | 14    | 22    |
| AltCPKs| Arabidopsis thaliana    | 6–7         | 1455–1941 | 484–646 | 21    | 28    |
| GmCPKs| Glycine max              | 7           | 1377–1788 | 459–596 | 21    | 28    |
| SlCPKs| Solanum lycopersicum    | 7–8         | 1290–1797 | 430–599 | 15    | 23    |
| ScCPKs| Solanum tuberosum       | 7–8         | 1530–2626 | 510–638 | 13    | 22    |
| TkCPKs| Taraxacum kokssghya     | 6–8         | 145–1779  | 483–592 | 22    | 28    |
| CnCPKs| Chrysanthemum nankingense| 6–9        | 1230–1815 | 409–604 | 11    | 22    |
| HaCPKs| Helianthus annuus       | 7–9         | 1470–1896 | 490–632 | 25    | 35    |
| LsCPKs| Lactuca sativa          | 6–7         | 1464–1800 | 488–600 | 20    | 28    |
| CcCPKs| Cynara cardunculus      | 7           | 1470–1824 | 489–607 | 18    | 25    |
| HbcCPKs| Hvezia brasiliensis    | 7           | 1110–1773 | 369–590 | 18    | 25    |

* All sequences used in this work are provided in Appendix A. Pal: palmitoylation; Myr: myristoylation. The palmitoylation and myristoylation sites were predicted using Terminator (https://bioweb.i2bc.paris-saclay.fr/terminator3/).

2.2. Phylogenetic Analysis of CPK Members

The five Asteraceae species CPKs, rubber tree CPKs, as well as five representative model plant CPK members were used to investigate the evolutionary relationship of CPKs in Asteraceae plants and rubber-producing plants. The result revealed that all CPK genes fell into four different groups, group I (yellow), group II (blue), group III (green) and group IV (pink) (Figure 1). No species-specific clades and rubber-producing-specific clades were identified. However, the amounts of CPKs in different groups of eleven species are different. Group I usually has the largest number and group IV contains the fewest number of CPKs.

2.3. Evolutionary Analyses of Duplicated Gene Pairs in Asteraceae Species

To explore the evolution of the CPK family in detail, we analyzed the duplication events of the Asteraceae CPK gene family. A number of 10, 11, 20, 13 and 7 duplicated paralogs were identified in Tks, L sativa, H. annuus, C. nankingense and C. cardunculus, respectively. The Ka/Ks ratio was calculated to assess the selection pressure of each duplicated paralog pairs. The results showed that most duplicated Asteraceae CPK paralogs are under purifying selection, except six duplicated gene pairs (HaCPK2/HaCPK23, HaCPK9/HaCPK19, HaCPK9/HaCPK28, CnCPK5/CnCPK28, CnCPK16/CnCPK31 and CnCPK18/CnCPK26), which are under positive selection (Figure 2, Tables S1–S5). It therefore appears that CPKs play critical roles during plant development, which requires highly conserved sequences. The paralogs under positive selection might have potential functional divergence which is involved in specific tissues, and development processes in H. annuus and C. nankingense after the emergence of Asteraceae.
Figure 1. Phylogenetic tree of CPK members in eleven species. The phylogenetic tree was constructed by ClustalX 2.0 and MEGA 7.0 software using the Neighbor-Joining (NJ) method with bootstrap replicates set to 1000. The CPKs from different species are shown in different colors. The constructed phylogenetic tree showed that CPKs are clustered into four different groups: group I (yellow), group II (blue), group III (green) and group IV (pink).

Figure 2. The Ka and Ks values of duplicated gene pairs in Asteraceae. The Ka and Ks values were calculated using Dnasp 4.0. Different colored shapes represent the Ka/Ks ratio of duplicated gene pairs from corresponding species.

We also addressed the question of whether these duplicated paralogs of Asteraceae CPKs are under an accelerated evolutionary rate. To this end, we assessed the Tajima relative rates of Asteraceae CPK paralogs. The TkCPK26/TkCPK27 and TkCPK27/TkCPK31 duplication pairs have prominently accelerate evolutionary rates (Table 2). The number of duplicated gene pairs from group III were more than that from the other three groups in the five Asteraceae species. Meanwhile, a total of 16 duplicated gene pairs (LsCPK19/LsCPK28, LsCPK20/LsCPK31, LsCPK11/LsCPK18, LsCPK15/LsCPK29, LsCPK16/LsCPK18, HaCPK33/HaCPK34, HaCPK33/HaCPK38, HaCPK10/HaCPK16, HaCPK19/HaCPK28, CnCPK14/CnCPK25, CnCPK29/CnCPK34, CnCPK16/CnCPK31, CnCPK20/CnCPK24, CnCPK2/CnCPK30, CnCPK1/CnCPK26, CnCPK18/CnCPK26) are under accelerated evolutionary rates in the other four Asteraceae species, suggesting that they potentially play specific roles. (Tables S6–S9).
Table 2. Tajima relative rate tests of CPK gene pairs in *Taraxacum koksaqhyza*.

| Testing Group               | Group | Mt  | M1 | M2 | χ²  | p   |
|-----------------------------|-------|-----|----|----|-----|-----|
| TkCPK7/TkCPK11 with SlCPK20 | II    | 408 | 0  | 0  | 0.00| 1.00000 |
| TkCPK9/TkCPK15 with SlCPK24 | III   | 357 | 0  | 0  | 0.00| 1.00000 |
| TkCPK25/TkCPK33 with SlCPK23| III   | 452 | 0  | 1  | 1.00| 0.31731 |
| TkCPK26/TkCPK31 with SlCPK8 | I     | 445 | 0  | 0  | 0.00| 1.00000 |
| TkCPK8/TkCPK21 with SlCPK25 | III   | 454 | 13 | 13 | 0.00| 1.00000 |
| TkCPK26/TkCPK27 with SlCPK8 | I     | 417 | 28 | 11 | 2.81| 0.09345 |
| TkCPK27/TkCPK31 with SlCPK8 | I     | 417 | 28 | 11 | 2.81| 0.09345 |
| TkCPK4/TkCPK19 with SlCPK26 | III   | 425 | 16 | 11 | 0.93| 0.33592 |
| TkCPK4/TkCPK32 with SlCPK26 | III   | 409 | 27 | 28 | 0.02| 0.87590 |

* The Tajima relative rate test was used to examine the equality of evolutionary rate between *Tk* paralogs; **Mt** is the sum of the identical sites in all three sequences tested; **M1** is the number of unique differences in the first paralog; **M2** is the number of unique differences in the second paralog; **p** if *p* < 0.05, the test rejects the equal substitution rates between the two duplicates and infers that one of the two duplicates has an accelerated evolutionary rate.

2.4. Syntenic Analysis of CPKs from Five Asteraceae Species

The syntenic analysis of CPK members from *Tk*, *L. sativa*, *H. annuus*, *C. nankingense*, *C. cardunculus* and *S. lycopersicum* was performed. The Circos program was used to visualize the syntenic relationship. A total of 10, 11, 20, 13 and 7 duplicated CPK pairs in *Tk*, *L. sativa*, *H. annuus*, *C. nankingense* and *C. cardunculus* were identified. The number of 10, 4, 7 and 1 TkCPKs from group I, group II, group III and group IV had syntenic relationships with CPKs from the other four *Asteraceae* species and tomato (Figure 3). Overall, there is a close CPK syntenic relationship among the five *Asteraceae* species, especially among different subgroups.

**Figure 3.** Duplication and syntenic analysis of CPKs in *Asteraceae* plants and *S. lycopersicum*. Chromosomes and scaffolds are shown in different colors and all CPK gene locations are indicated.
The syntenic relationship of CPKs in four different groups are connected by different colored lines: group I (yellow), Group II (red), group III (magenta) and Group IV (blue).

2.5. The CPKs in Group II and Group III Are Expanded

To investigate the evolution of the CPK gene family, 39 GmCPKs in Leguminosae were chosen to represent CPKs in eurosids I. In addition, 34 AtCPKs in Brassicaceae represented CPKs in eurosids II, and 29 Solanum lycopersicum CPKs, 26 Solanum tuberosum CPKs, 28 Nicotiana tabacum CPKs, 31 Capsicum annuum CPKs and 40 Ipomoea nil CPKs in Solanaceae represented CPKs from euasterids I. Five Asteraceae species stand for CPKs from euasterids II and 30 OsCPKs in Gramineae represented CPKs in Monocotyledon. By comparing the CPK number in all mentioned species, we found that the CPK gene family has the largest number in group I and the smallest numbers in group IV. The AtCPKs and GmCPKs in rosids are significantly expanded in group II, while CPKs in Asteraceae are expanded in group III (Figure 4), indicating the potential functional divergence of expanded CPKs in group II and group III in rosids and Asteraceae, respectively. Notably, rubber-producing Asteraceae plants have the largest group III members (12, 11 and 11 for Tks, Ha and Ls), implying members of group III might be potentially involved in NR-related metabolism processes.

Figure 4. Expansion of CPKs in different groups. Thirteen species are ordered by plant classification system. The number of CPKs of four distinct groups from different species are indicated in the colored boxes. Red and green represent a large or a small number of CPK members, respectively.

2.6. Gene Structure and Motif Distribution of CPKs

Gene structure divergence plays considerable roles in gene family evolution and can be used to assess phylogenetic relationships [33,34]. To further investigate the expansion mechanism in group III, maps of exon–intron structure and motif distribution were constructed based on coding DNA sequences as well as protein sequences of group III CPKs from five model plants and five Asteraceae species. The result displayed a very similar exon–intron structures of ten species in group III. The first exon in most CPK members was the longest, followed by several shorter exons. Meanwhile, the gene structure and motif distributions showed similar patterns in other three groups (Figure 5, Figures S1–S3).
Figure 5. Gene structure and conserved motif distribution of group III CPKs from five model plants and five Asteraceae species. The phylogenetic tree-view is on the left panel. Exon–intron distribution is in the middle, in which black lines and green boxes represent introns and exons, respectively. The motif distribution is on the right panel, in which rectangles with different colors represent different conserved motifs.

2.7. Motif Sequence Analysis of CPKs in Group III from Asteraceae

The CPK members of Asteraceae were significantly expanded in group III. The exon–intron structure and motif distribution analysis showed that all CPKs in group III from Asteraceae were conservative (Figure 5). To further explore the tendency of the CPKs’ expansion in group III, the detailed information of conserved motifs (amino acid sequences) were analyzed (Figure 6). The result showed that the DLK motif and auto-inhibitory domain of all four groups were highly conserved. However, lower conservativeness was observed in EF-hand 1 to EF-hand 3 of group III CPKs from five Asteraceae species. In Asteraceae, the “D1-X-D3-X-S5” regions of group III in EF-hand 1 to EF-hand 3 have an obviously lower convergence than that of group I, II and IV. (framed in Figure 6). The Ca²⁺-binding sites of the EF-hands were reported to be D1-D3-S5-E12 (EF-hand 1), D1-D3-S5-E12 (EF-hand 2), D1-D3-S5-E12 (EF-hand 3), and D1-D3-D5-E12 (EF-hand 4) [35].
It seems that group III CPKs from Asteraceae might have lower calcium binding capacity than other groups since the amino acids of the EF loop region participating in Ca\(^{2+}\)-binding.

Figure 6. Comparison of sequence logos for conserved motifs of CPKs from the five Asteraceae species. Sequence logos of the consensus motifs were created using MEME online software. The height of each letter represents the frequency of amino acids at corresponding position. The red star means all CPK members have the exact same amino acid in corresponding site. Black frames showed the amino acid sites with much lower conservation in EF-hand 1, 2 and 3 of group III.

3. Discussion

3.1. Identification and Characteristics of CPKs in Asteraceae Species

Genome-wide identification of the CPK family has been conducted in various higher plants [28–32,36–39]. A total of 34, 34, 40, 34 and 30 novel CPKs and 10, 11, 20, 13 and 7 duplication gene pairs in Tks, L sativa, H. annuus, C. nankingense and C. cardunculus were identified. (Table 2, Tables S1–S5). Four species have a similar number of CPKs to that of Arabidopsis and rice, expect H. annuus, in which the CPK number significantly expanded, which may be ascribed to a much more complex evolutionary history experienced by sunflower with a lineage-specific whole-genome duplication (WGD) event around 29 million years ago [14]. In the five Asteraceae species studied in this study, an additional WGD event had been identified only in H. annuus, but not in the other four species [14–18].

Gene structure analysis of CPKs showed that the first exon in most CPKs was the longest one, followed by several shorter exons. The exon number in the four sub-groups were different, CPKs in group IV had more but shorter exons than groups I–III. The exon–intron patterns were similar between CPKs belonging to the same evolutionary groups (Table 1, Figure 5). Further, duplicated CPK gene pairs had highly conserved exon–intron patterns, which may also impact on the functional similarities and/or redundancy between these duplicated genes.

The CPK sequences among all higher plants are highly conserved, particularly in the protein kinase domain, the auto-inhibitory domain and the four EF-hand domains [40]. The CPKs of five Asteraceae species in this study are also highly conserved. Gene structure and motif distribution analyses of CPKs showed that the members in the same group shared similar distribution patterns of exon–introns and motifs (Figure 5 and Figures S1–S3).
Moreover, the Ka/Ks ratio among paralogs in five Asteraceae species demonstrated that evolutionary pressure for these sequences was maintained as most Ka/Ks ratios are less than 1 (Figure 2), indicating that these CPKs are under purifying selection.

### 3.2. Phylogenetic Analysis and Group-Specific Expansion of CPKs in Asteraceae Species

The CPKs in group IV appeared to diverge from the common ancestor with algae; group III formed a clade separate from groups I and II, while the split between group I and II appeared to be the most recent evolutionary event. The phylogenetic tree of CPK members in eleven species revealed that the group IV have the longest main branch followed by group III, and the branch of group II is the shortest (Figure 1). Furthermore, the exon–intron numbers and distribution in group IV were also different from the other three groups (Table 1, Figure 5 and Figures S1–S3), supporting the hypothesis that group IV CPKs form a separate clade of earlier lineage [32].

The CPK gene family has expanded greatly from four genes in the land plant ancestor, and less than 11 genes in green algae to approximately 30–40 members among angiosperms. Our phylogenetic analysis provides insights regarding the evolutionary relationship and group-specific expansion of CPKs from Asteraceae and rosids. CPKs in group II and group III are significantly expanded in rosids and Asteraceae, respectively (Figure 4). Gene replication contributes to the expansion of gene families. In these five Asteraceae species, the numbers of paralogs in group III were much higher than in other groups, indicating that gene replication was the main reason of the group III CPKs' expansion.

The AtCPK10 and AtCPK30 play a central role in regulating primary nitrate responses and controlling of primary transcription by the RNA sequencing [41], suggesting that TkCPK25/TkCPK33, located in the same phylogenetic tree branch with AtCPK10 and AtCPK30, might also take part in the primary transcription regulation. In addition, TkCPK4/TkCPK19 and TkCPK4/TkCPK32 may participate in drought stress regulation since the ortholog gene AtCPK8 functions in ABA-mediated stomatal movement in response to drought stress through the regulation of catalase 3 [42]. AtCPK24 could negatively regulate pollen tube growth by inhibiting K\(^+\) inward currents [43], indicating that TkCPK9, TkCPK15 and TkCPK16 might also be involved in the development of pollen tube. Notably, within the five Asteraceae species, more group III members were observed in the three rubber-producing plants (12 for Tks, 11 for H. annuus and 11 for L. sativa), compared to non-rubber species (Figure 4). For rubber-producing plants, the HbCPKs also show a slightly expansion in group III (nine members) compared with other groups [32], indicating that there might be some potential roles for group III CPKs in NR biosynthesis.

Previous research investigated the sequence degeneration of group III CPK Ca\(^{2+}\)-binding sites, showing that five AtCPKs (CPK 7, 8, 10, 13 and 32) have lower or no calcium sensitivity [44]. All these weak CPKs carry one or two altered EF-hand motif(s), suggesting that the degeneration of the EF-hand motifs can greatly influence the calcium loading. AtCPK13, a member of group III, inhibits stomatal opening under light-induced conditions [45], indicating their orthologs TkCPK8/TkCPK21 may also be involved in a similar pathway. The expanded group III CPKs from Asteraceae exhibited less conservative in EF-hands of “D-X-D-X-S” region than other three groups (framed in Figure 6). Unlike EF-hands 1 to 3, EF-hands 4 of Asteraceae group III CPKs still showed high conservation, implying their importance to the CPK Ca\(^{2+}\)-binding capacity.

### 4. Materials and Methods

#### 4.1. Identification and Characteristics of CPK Members in Five Representative Asteraceae Species

The protein sequences of CPKs from Arabidopsis and rice served as a query sequence to perform the local BLASTP program for identifying CPK members in Tks, L. sativa, H. annuus, C. nankingense and C. cardunculus (e-value < 1 × 10\(^{-5}\)).

InterProScan (http://www.ebi.ac.uk/interpro/) was used to determine whether all putative CPK candidates contain the protein kinase domain (PF00069) and EF-hand_7 domain (PF13499). CPK members from five representative model species and the rubber tree are derived from previous studies [5,28–32].
Plants 2020, 9, 32

The myristoylation site and palmitoylation site were predicted by software TermiNator (https://bioweb.i2bc.paris-saclay.fr/terminator3/).

4.2. Phylogenetic Tree Construction

Amino acid sequences of CPKs from A. thaliana, O. sativa, L. sativa, H. annuus, G. max, T. koksaghyz, C. nankingense, S. tuberosum, S. lycopersicum, C. cardunculus and H. brasiliensis were aligned by ClustalX 2.0. The phylogenetic tree was constructed using MEGA 7.0 using Neighbor-Joining method, 1000 bootstrap replications [46].

4.3. Duplication Event and Syntenic Analysis.

Paralogs of CPKs from five Asteraceae species were determined by multiple sequence alignment with the amino acid identification > 80%. The Ka/Ks ratios for these CPK paralogs were calculated to evaluate the selection pressure; the ratio > 1, < 1, or = 1 indicates positive, negative or neutral evolution, respectively. The Ka/Ks ratios of these paralogs were calculated using Dnasp 4.0 software [47]. Tajima relative rate tests were detected by MEGA 7.0 using the amino acid sequences of the duplicated CPK pairs [48]. The result of local BLASTp program (with an E-value setting of $1 \times 10^{-5}$) and the sorted GFF profiles (with four columns, the first column is chromosome name, the second column is gene name, the third column is gene starting position and the last column is gene ending position) were then submitted to the MCScan program to identify the syntenic relationships of paralogs and/or orthologs of CPKs among six species [49]. Circos program [50] was used for visualizing the syntenic results.

4.4. Gene Structure and Motif Distribution Analysis

The gene structures of CPK members from ten species were constructed using TBtools JRE1.6 [51] based on the genomic sequence and coding DNA sequences corresponding to each predicted gene. The conserved motifs for all CPK protein sequences and conserved CPK sequence motif logos of five Asteraceae species were detected by Multiple Expectation Maximization for Motif Elicitation (MEME) online tools (http://meme.sdsc.edu/meme/intro.html).

5. Conclusions

In summary, our study provides a comprehensive evolutionary and systematical analysis of CPK members in five representative Asteraceae species, showing that for the representative rubber-producing Asteraceae plant Tks, duplicated gene pairs were under purifying selection pressure and two TkCPK duplication paralogs had an accelerated evolutionary rate. By comparing the CPK numbers in four groups, we found that CPKs in group II and group III were significantly expanded in rosids and rubber-producing Asteraceae plants, indicating that potential functional divergence of expanded CPKs in group II and group III in rosids and rubber-producing Asteraceae plants, respectively. Further gene structure and motif distribution analyses in group III revealed that the exon–intron and motif distribution were similar and conserved. Detailed conserved motif logos analysis revealed that CPKs in group III of Asteraceae species have lower amino acid conservation in EF-hand I to III, indicating that they might have lower calcium binding ability than the other three groups. Our data provide a systematical evolutionary investigation of the CPKs in five representative Asteraceae species, suggesting sub-family specific expansion of CPKs might be related to natural rubber producing.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Gene structure and conserved motif distribution of CPKs from group I, Figure S2: Gene structure and conserved motif distribution of CPKs from group II, Figure S3: Gene structure and conserved motif distribution of CPKs from group IV, Table S1: The Ka/Ks ratios for duplicated CPK genes in T. koksaghyz, Table S2: The Ka/Ks ratios for duplicated CPK genes in H. annuus, Table S3: The Ka/Ks ratios for duplicated CPK genes in L. sativa, Table S4: The Ka/Ks ratios for duplicated CPK genes in C. cardunculus, Table S5: The Ka/Ks ratios for duplicated CPK genes in C. nankingense, Table S6: Tajima relative rate tests of CPK gene pairs in H. annuus, Table S7: Tajima relative
rate tests of CPK gene pairs in *L. sativa*, Table S8: Tajima relative rate tests of CPK gene pairs in *C. cardunculus*, Table S9: Tajima relative rate tests of CPK gene pairs in *C. nankingense*.

**Author Contributions:** Conceptualization L.Z. and X.J.; methodology, L.Z. and B.Z.; software, B.Z.; validation, L.Z., B.Z. and W.S.; formal analysis, B.Z.; investigation, W.S.; resources, L.Z.; data curation, L.Z.; writing—original draft preparation, L.Z.; writing—review and editing, H.L. and X.J.; visualization, H.L.; supervision, H.L.; project administration, X.J.; funding acquisition, X.J. and H.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Natural Science Foundation of Hainan Province, grant number 319MS046, National Natural Science Foundation of China, grant number 31860068, International Scientific and Technological Cooperation project by Shihezi University, grant number GJHZ201708, Science and Technology Project of Bingtuan, grant number 2018AB012.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Abbreviations**

| Abbreviation | Description |
|--------------|-------------|
| CPK          | calcium-dependent protein kinases |
| CaM          | calmodulins |
| CaML         | CaM-like proteins |
| CBL          | calcineurin B-like proteins |
| NR           | natural rubber |
| cPTs         | cis-prenyltransferases |
| SRPPs        | small rubber particle proteins |
| SNP          | single nucleotide polymorphism |
| Pal          | palmitoylation |
| Myr          | myristoylation |
| CDS          | coding sequence |
| Ka           | nonsynonymous substitution rate |
| Ks           | synonymous substitutions rate |
| NJ           | Neighbour-Joining |
| MEME         | Multiple Expectation Maximization for Motif Elicitation |
| WGD          | whole-genome duplication |

**Appendix A**

Appendix A provides all CPK sequences used in this study.

**References**

1. White, P.; Broadley, M. Calcium in plants. *Ann. Bot.* **2003**, *92*, 487–511.
2. Luan, S. The CBL-CIPK network in plant calcium signaling. *Trends Plant Sci.* **2009**, *14*, 37–42.
3. Ludwig, A.A.; Romeis, T.; Jones, J.D. CDPK-mediated signalling pathways: Specificity and cross-talk. *J. Exp. Bot.* **2004**, *55*, 181–188.
4. Harper, J.F.; Sussman, M.R.; Schaller, G.E.; Putnam-Evans, C.; Charbonneau, H.; Harman, A.C. A calcium-dependent protein kinase with a regulatory domain similar to calmodulin. *Science* **1991**, *252*, 951–954.
5. Cheng, S.H.; Willmann, M.R.; Chen, H.C.; Sheen, J. Calcium signalling through protein kinases. The *Arabidopsis* calcium-dependent protein kinase gene family. *Plant Physiol.* **2002**, *129*, 469–485.
6. Saito, S.; Hamamoto, S.; Moriga, K.; Matsuura, A.; Sato, Y.; Muto, J.; Noguchi, H.; Yamauchi, S.; Tozawa, Y.; Ueda, M.; et al. N-myristoylation and S-acylation are common modifications of Ca\(^{2+}\)-regulated *Arabidopsis* kinases and are required for activation of the SLAC1 anion channel. *New Phytol.* **2018**, *218*, 1504–1521.
7. Yip Delormel, T.; Boudsocq, M. Properties and functions of calcium-dependent protein kinases and their relatives in *Arabidopsis thaliana*. *New Phytol.* **2019**, *224*, 585–604.
8. Boudsocq, M.; Droillard, M.J.; Regad, L.; Laurière, C. Characterization of *Arabidopsis* calcium-dependent protein kinases: Activated or not by calcium? *Biochem. J.* **2012**, *447*, 291–299.
9. Franz, S.; Ehler, B.; Liese, A.; Kurth, J.; Cazalé, A.C.; Romeis, T. Calcium-dependent protein kinase CPK21 functions in abiotic stress response in *Arabidopsis thaliana*. *mol. Plant* **2011**, *4*, 83–96.
10. Tähtiharju, S.; Rijpkema, A.S.; Vetterli, A.; Albert, V.A.; Teeri, T.H.; Elomaa, P. Evolution and diversification of the CYC/TB1 gene family in Asteraceae—A comparative study in Gerbera (Mutisieae) and sunflower (Helianthae). Mol. Biol. Evol. 2012, 29, 1155–1166.

11. Barker, M.S.; Kane, N.C.; Matvienko, M.; Kozik, A.; Michelmore, R.W.; Knapp, S.J.; Rieseberg, L.H. Multiple paleopolyploidizations during the evolution of the Compositae reveal parallel patterns of duplicate gene retention after millions of years. Mol. Biol. Evol. 2008, 25, 2445–2455.

12. Funk, V.A.; Bayer, R.J.; Keeley, S.; Chan, R.; Watson, L.; Gemeinholzer, B.; Schilling, E.; Panero, J.L.; Baldwin, B.G.; Garcia-jacas, N.; et al. Everywhere but Antarctica: Using a supertree to understand the diversity and distribution of the Compositae. Biol. Skr. 2005, 55, 343–374.

13. Judd, W.S.; Campbell, E.A. Kellogg. Plant systematics: A phylogenetic approach. Cladistics Int. J. Willi Hennig Soc. 2010, 24, 848–850.

14. Badouin, H.; Gouzy, J.; Grassa, C.J.; Murat, F.; Staton, S.E.; Cottret, L.; Lelandais-Brière, C.; Owens, G.L.; Carrière, S.; Mayjonade, B.; et al. The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. Nature 2017, 546, 148–152.

15. Reyes-Chin-Wo, S.; Wang, Z.; Yang, X.; Kozik, A.; Arikitt, S.; Song, C.; Xia, L.; Froenicke, L.; Lavelle, D.O.; Truco, M.J.; et al. Genome assembly with in vitro proximity ligation data and whole-genome triplication in lettuce. Nat. Commun. 2017, 8, 14953.

16. Lin, T.; Xu, X.; Ruan, J.; Liu, S.; Wu, S.; Shao, X.; Wang, X.; Gan, L.; Qin, B.; Yang, Y.; et al. Genome analysis of Taraxacum kok-saghyz rodin provides new insights into rubber biosynthesis. Natl. Sci. Rev. 2018, 5, 78–87.

17. Scaglione, D.; Reyes-Chin-Wo, S.; Acquandio, A.; Froenicke, L.; Portis, E.; Beitel, C.; Tirone, M.; Mauro, R.; Lo, M.A.; Maurumicalle, G.; et al. The genome sequence of the outbreeding globe artichoke constructed de novo incorporating a phase-aware low-pass sequencing strategy of F. progeny. Sci. Rep. 2016, 6, 19427.

18. Song, C.; Liu, Y.; Song, A.; Dong, G.; Zhao, H.; Sun, W.; Ramakrishnan, S.; Wang, Y.; Wang, S.; Li, T.; et al. The Chrysanthemum nankingense genome provides insights into the evolution and diversification of chrysanthemum flowers and medicinal traits. Mol Plant. 2018, 11, 1482–1491.

19. Tang, C.; Yang, M.; Fang, Y.; Luo, Y.; Gao, S.; Xiao, X.; An, Z.; Zhou, B.; Zhang, B.; Tan, X.; et al. The rubber tree genome reveals new insights into rubber production and species adaptation. Nat. Plants 2016, 2, 16073.

20. Stipanovic, R.; O’Brien, D.; Rogers, C.; Hanlon, K. Natural rubber from sunflower. J. Agric. Food Chem. 1980, 28, 1322–1323.

21. Schmidt, T.; Hillebrand, A.; Wurbs, D.; Wahler, D.; Lenders, M.; Gronover, C.S.; Prüfer, D. Molecular cloning and characterization of rubber biosynthetic genes from Taraxacum kok-saghyz. Plant Mol. Biol. Rep. 2010, 28, 277–284.

22. Collins-Silva, J.; Nural, A.T.; Skaggs, A.; Scott, D.; Hathwaik, U.; Woolsey, R.; Schegg, K.; McMahan, C.; Whalen, M.; Cornish, K.; et al. Altered levels of the Taraxacum kok-saghyz (Russian dandelion) small rubber particle protein, TkSRP3, result in qualitative and quantitative changes in rubber metabolism. Phytochemistry 2012, 79, 46–56.

23. Qu, Y.; Chakrabarty, R.; Tran, H.T.; Kwon, E.J.; Kwon, M.; Nguyen, T.D.; Ro, D.K. A lettuce (Lactuca sativa) homolog of human Nogo-B receptor interacts with cis-prenyltransferase and is necessary for natural rubber biosynthesis. J. Biol. Chem. 2015, 290, 1898–1914.

24. Epping, J.; Deenen, N.V.; Niephaus, E.; Stolze, A.; Fricke, J.; Huber, C.; Eisenreich, W.; Twyman, R.M.; Prüfer, D.; Gronover, C.S. A rubber transferase activator is necessary for natural rubber biosynthesis in dandelion. Nat. Plants 2015, 1, 15048. doi:10.1038/nplants.2015.48.

25. Luo, Z.; Iaffaldano, B.J.; Zhuang, X.; Fresno, Ramirez, J.; Cornish, K. Analysis of the first Taraxacum kok-saghyz transcriptome reveals potential rubber yield related SNPs. Sci. Rep. 2017, 7, 9939.

26. Zhu, L.; Jin, X.; Xie, Q.; Yao, Q.; Wang, X.; Li, H. Calcium-dependent protein kinase family genes involved in ethylene-induced natural rubber production in different Hevea brasiliensis cultivars. Int. J. Mol. Sci. 2018, 19, 947.

27. Jin, X.; Zhu, L.; Yao, Q.; Meng, X.; Ding, G.; Wang, D.; Xie, Q.; Tong, Z.; Tao, C.; Yu, L.; et al. Expression profiling of mitogen-activated protein kinase genes reveals their evolutionary and functional diversity in different rubber tree (Hevea brasiliensis) cultivars. Genes 2017, 8, 261.

28. Ray, S.; Agarwal, P.; Arora, R.; Kapoor, S.; Tyagi, A.K. 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. Genom. 2007, 278, 493–505.

29. Liu, H.; Che, Z.; Zeng, X.; Zhou, X.; Sitoe, H.M.; Wang, H.; Yu, D. Genome-wide analysis of calcium-dependent protein kinases and their expression patterns in response to herbivore and wounding stresses.
30. Hu, Z.; Lv, X.; Xia, X.; Zhou, J.; Shi, K.; Yu, J.; Zhou, Y. Genome-wide identification and expression analysis of calcium-dependent protein kinase in tomato. *Front. Plant Sci.* 2016, 7, 469.

31. Fantino, E.; Segretin, M.E.; Santin, F.; Mirkin, F.G.; Ulloa, R.M. Analysis of the potato calcium-dependent protein kinase family and characterization of StCDPK7, a member induced upon infection with *Phytophthora infestans*. *Plant Cell Rep.* 2017, 36, 1137–1157.

32. Xiao, X.H.; Yang, M.; Sui, J.L.; Qi, J.Y.; Fang, Y.J.; Hu, S.N.; Tang, C.R. The calcium-dependent protein kinase (CDPK) and CDPK-related kinase gene families in *Hevea brasiliensis*-comparison with five other plant species in structure, evolution, and expression. *FEBs Open* 2016, 7, 4–24.

33. Boudet, N.; Aubourg, S.; Toffano-Nioche, C.; Kreis, M.; Lechamry, A. Evolution of intron/exon structure of DEAD helicase family genes in *Arabidopsis, Caenorhabditis*, and *Drosophila*. *Genome Res.* 2001, 11, 2101–2114.

34. Wei, K.; Chen, J.; Wang, Y.; Chen, Y.; Chen, S.; Lin, Y.; Pan, S.; Zhong, X.; Xie, D. Genome-wide analysis of bZIP-encoding genes in maize. *DNA Res.* 2012, 19, 463–476.

35. Shi, S.; Li, S.; Asim, M.; Mao, J.; Xu, D.; Ullah, Z.; Liu, G.; Wang, Q.; Liu, H. The Arabidopsis calcium-dependent protein kinase genes (CDPKs) and their roles in plant growth regulation and abiotic stress responses. *Int. J. Mol. Sci.* 2018, 19, 1900.

36. Kong, X.; Lv, W.; Jiang, S.; Zhang, D.; Cai, G.; Pan, J.; Li, D. Genome-wide identification and expression analysis of calcium-dependent protein kinase gene family in maize. *BMC Genom.* 2013, 14, 433.

37. Zhang, K.; Han, Y.; Zhao, F.; Hu, Y.; Gao, Y.; Ma, Y.; Zheng, Y.; Wang, Y.; Wen, Y. Genome-wide identification and expression analysis of the CDPK gene family in grape, Vitis spp. *BMC Plant Biol.* 2015, 15, 164.

38. Xu, X.; Liu, M.; Lu, L.; He, M.; Qu, W.; Xu, Q.; Qi, X.; Chen, X. Genome-wide analysis and expression of the calcium-dependent protein kinase gene family in cucumber. *Mol. Genet. Genom.* 2015, 290, 1403–1414.

39. Liu, W.; Li, W.; He, Q.; Daud, M.K.; Chen, J.; Zhu, S. Genome-wide survey and expression analysis of calcium-dependent protein kinase in *Gossypium raimondii*. *PLoS ONE* 2014, 9, e98189.

40. Valmonte, G.R.; Arthur, K.; Higgins, C.M.; MacDiarmid, R.M. Calcium-dependent protein kinases in plants: Evolution, expression and function. *Plant Cell Physiol.* 2014, 55, 551–569.

41. Liu, K.H.; Niu, Y.; Konishi, M.; Wu, Y.; Du, H.; Sun, C.H.; Li, L.; Boudsocq, M.; McCormack, M.; Maekawa, S.; et al. Discovery of nitrate-CPK-NLP signalling in central nutrient-growth networks. *Nature* 2017, 18, 311–316.

42. Zou, J.J.; Li, X.D.; Ratnashekera, D.; Wang, C.; Liu, W.X.; Song, L.F.; Zhang, W.Z.; Wu, W.H. *Arabidopsis* calcium-dependent protein kinase and catalase 3 function in abscisic acid-mediated signaling and H2O2 homeostasis in stomatal guard cells under drought stress. *Plant Cell* 2015, 27, 1445–1460.

43. Zhao, L.N.; Shen, L.K.; Zhang, W.Z.; Zhang, W.; Wang, Y.; Wu, W.H. Ca2+-dependent protein kinase11 and 24 modulate the activity of the inward rectifying K+ channels in *Arabidopsis* pollen tubes. *Plant Cell* 2013, 25, 649–661.

44. Boudsocq, M.; Sheen, J. CDPKs in immune and stress signaling. *Trends Plant Sci.* 2013, 18, 30–40.

45. Ronzier, E.; Corratgé-Faillle, C.; Sanchez, F.; Prado, K.; Briere, C.; Leonhardt, N.; Thibaud, J.B.; Xiong, T.C. CPK13, a noncanonical Ca2+-dependent protein kinase, specifically inhibits KAT2 and KAT1 shaker K+ channels and reduces stomatal opening. *Plant Physiol.* 2014, 166, 314–326.

46. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 2016, 33, 1870–1874.

47. Librado, P.; Rozas, J. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 2009, 25, 1451–1452.

48. Tajima, F. Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics* 1993, 135, 599–607.

49. Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.H.; Jin, H.; Marker, B.; Guo, H.; et al. MCScanX: A toolkit for detection and evolutionary analysis of gene synten and collinearity. *Nucleic Acids Res.* 2012, 40, e49.

50. Krzywinski, M.; Schein, J.; Birol, I.; Connors, J.; Gascoyne, R.; Horsman, D.; Jones, S.J.; Marra, M.A. Circos: An information aesthetic for comparative genomics. *Genome Res.* 2009, 19, 1639–1645.