Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Genome-Wide Characterization of CDPK Gene Family in Apple (Malus Domestica) and Its Transcriptional Expression During Apple Fruit Development

Yu Dai
Anhui Agricultural University
Xinxin Zhang
Anhui Agricultural University
Juan Tian
Anhui Agricultural University
Liufang Zhu
Anhui Agricultural University
Dahui Li (dahui2@126.com)
Anhui Agricultural University

Research Article

Keywords: Apple, CDPK, Fruit development, Gene family, Transcriptional expression

Posted Date: April 30th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-445130/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Abstract

Background: Calcium-dependent protein kinases (CDPKs) play important roles both in developments and response to stresses, via mediating Ca^{2+} signal transduction in plants. To characterize the CDPKs in apple (*Malus domestica*), the apple *CDPK* gene family, together with those from pear (*Pyrus bretschneideri*), peach (*Prunus persica*), strawberry (*Fragaria vesca*), and and *Arabidopsis thaliana*, were analyzed at the genome-wide level in the present study.

Results: A total of 116 *CDPK*s, consisting of 24 *MdCDPK*s, 28 *PbCDPK*s, 16 *PpCDPK*s, 14 *FvCDPK*s, and 34 *AtCDPK*s, was identified from apple, pear, peach, strawberry, and Arabidopsis, respectively. An integrated analysis of these *CDPK*s was performed on their chromosomal distribution, phylogenetic and collinearity relationships, characteristics of gene structures and conserved motifs. As a result, the *CDPK* gene family members were showed to be highly conserved both at their kinase and EF-hand domains. Among 209 gene-pairs with interspecies collinearity, there existed 22, 36, 21, and 25 ones between *MdCDPK*s and other *CDPK*s in Arabidopsis, pear, strawberry, and peach, respectively. And the evaluated *Ka*/*Ks* ratios were less than 1 between the *CDPK* gene pairs with collinearity relationships. Transcriptomic analysis demonstrated that among 24 members of the apple *CDPK* gene family, two up-regulatory ones (*HF05266* and *HF09216*) and two down-regulatory ones (*HF05471* and *HF15429*), were differentially expressed with significance between the apple fruit developmental stage S4 (mature) and other stages (early growing-S1, mid growing-S2, and late growing-S3), respectively.

Conclusions: The whole genome duplication and subsequent purifying selection, might have played an important role in the *CDPK* gene expansion, leading to structural and functional novelty during evolution of the species lineages. In many cases, the *MdCDPK* genes within a phylogenetic group could show the different expression patterns at the transcriptional level, suggesting that these *MdCDPK*s have undergone genetic variant events and potential functional diversification. Some of *MdCDPK*s with significantly differential expression, were indicated their particular functions at the specific stages of apple fruit development.

Background

Plants are capable to regulate their physiological activities in response to various internal and outer signals, relying on signal transduction [1]. Calcium has been known as an important secondary messenger during the specific signaling pathways. The changes in cytosolic free Ca^{2+} elicited by certain stimuli, are coupled to peculiar cellular responses [2, 3]. Intracellular Ca^{2+} signaling may coordinate with other signaling pathways, generating a crosstalk network [4].

Following the elicitation of Ca^{2+} signature, calcium-binding proteins, for instance, calmodulins, calcineurin B-like proteins, and calcium-dependent protein kinases (CDPKs), play roes in signal transduction [2, 5]. Among these Ca^{2+} sensors identified in plants, CDPKs represent a type of protein characterized with the domains of protein kinase at their N-termini and EF-hand calcium-binding sites at their C-termini [6-8]. Thereby, upon direct binding to calcium, CDPKs are able to activate their calcium-stimulated kinase

activities, independence of the activation process mediated by calmodulins. In the case of lacking Ca^{2+} signature, CDPK kinase activities are auto-inhibited by a special sequence, namely junction domain, between the CDPK kinase and EF-hand domains [6-8]. The highly variable N-terminus of some CDPK, containing specific sites for myristoylation or palmitoylation, have been reported as membrane anchors [7, 9].

Ca^{2+} signals play versatile roles in regulating a various growth and developmental processes [10, 11]. It is therefore likely that CDPKs, as a class of Ca^{2+} sensors, have functions involved in the specific processes of plant development [8]. Some such cases associated with CDPKs, have been reported on embryogenesis, seed development and germination [12, 13], early stages of potato tuberization [14], pollen tube growth [15], and shoot growth [16].

Furthermore, CDPKs have been demonstrated to mediate adaptive regulations in response to a variety of abiotic and biotic stresses, such as cold, high salinity, drought, wounding, and pathogen infection [8, 17]. Transcriptional upregulation of *CDPK*s has been identified in a variety of species encountered by abiotic stresses [7, 17-19]. Arabidopsis CDPK10, and tobacco NtCDPK2 are involved in modulating osmotic potential [20]. OsCDPK7, a rice CDPK, has an important role in the tolerance to both cold and salt stress. The transgenic plants with overexpressed levels of *OsCDPK7*, showed an enhanced tolerance to cold, drought and salt stresses [21]. A set of the cotton CDPKs (GhCPK8, GhCPK38, GhCPK54, and GhCPK55) could participate in the early signaling events in cotton responses to salt stress [19].

With the completion of genome sequencing, characterization of *CDPK* gene family at the genome-wide level, has been carried out in many a plants species, such as Arabidopsis [7], maize [22], barley [23], and upland cotton [19]. The identified *CDPK* families in these species are composed of varied numbers of members, inferring their diverse functions during evolution.

Apple (*Malusdomestica* Borkh.) is one of the major fruit crops produced in the world. Its genome has been sequenced in 2017 [24]. Thereafter, a new version data of apple genome (HANFU) was released in 2019 (https://github.com/moold/Genome-data-of-Hanfu-apple). These data provide a better likelihood to excavate the *CDPK* gene family and should facilitate the elucidation of CDPK properties and functions in apples. In the present research, a comprehensive analysis of evolution and function of apple *CDPK*s was carried out at the whole-genome level. In addition to apple *CDPK*s, other *CDPK*s in three species of the *Rosaceae* family, including pear (*Pyrus bretschneideri*), strawberry (*Fragaria vesca*), and peach (*Prunus persica*), together with the model plant Arabidopsis (*A. thaliana*), were retrieved from the individual plant species with available genome data. The phylogenetic, gene structures and protein motifs of the identified *CDPK* family members, accompanied by the collinearity analysis on these genes, provided some clues to the evolutionary relationships among these *CDPK*s. Furthermore, to examine the apple *CDPK*s' involvement in the development of apple fruits, the transcriptomic data from RNA-seq analysis on two apple strains at the different stages of fruit development [25], were re-quantified to address this question. The results may lead to a primary understanding of the apple *CDPK*s with both redundant and distinct functions, and further investigating the function of calcium signaling mediated by the specific CDPK in regulating apple fruit development.

Methods

CDPK gene identification

The genome data of *A. thaliana*, *F. vesca*, and *P. persica*, were download from the database Phytozome (https://phytozome.jgi.doe.gov), while those of *M. domestica* from the online web (https://github.com/moold/Genome-data-of-Hanfu-apple) and *P. bretschneideri* from the Pear Genome Project (http://peargenome.njau.edu.cn). The *CDPK* gene family members were identified from the above mentioned five species. For CDPK identification, the CDPK-specific Hidden Markov Models (HMMs), Pkinase (PF00069) and EF-hand_7 (PF13499), were retrieved from the Pfam database (http://pfam.xfam.org/), using the Arabidopsis CDPK1 (AT5G04870) protein sequence [7, 19]. These HMMs were subsequently served as queries for scanning the genome data by the BLASTP program with an E-value less than 1e-5, respectively. The resultant sequences were further validated by their CDPK-specific motifs from the Pfam database.

Phylogenetic tree construction

For multiple sequence alignment based on amino acid sequences, all of the identified CDPK proteins with the full-length sequences or only the conserved domains (i.e. Pkinase and EF-hand_7), were aligned via the online program Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). The alignment result from the CDPK full-length sequences was further applied to the construction of phylogenetic tree with the Maximum Likelihood method in MEGA X.

Analysis of CDPK molecular characteristics, gene structures, conserved protein motifs

The primary molecular characteristics, such as amino acid length, isoelectric point (IP) and molecular weight, were analyzed by the online program ExPASy (http://www.expasy.ch/tools). Chromosomal localization of *CDPK* genes were visualized on the program package MapInspect (http://www.plantbreeding.wur.nl/). Analysis of *CDPK* gene structures was performed using the GSDS server (http://gsds.cbi.pku.edu.cn/). Analysis of the conserved motifs among the CDPKs was carried out using the program MEME (http://meme-suite.org/tools/meme) with default parameters.

Collinearity analysis of *CDPK* genes

For collinearity analysis, the local databases of proteins were built and blasted using the program package MCScanX [34]. The criterion for collinearity relationships was referred to the previous method [35]. For micro-synteny analysis between the *CDPK* gene pairs, flanking positions (i.e. a frame with 0.1-million base pairs of genomic sequences) located at both up- and downstream of the individual target genes, were used in a Blast comparison. *Ka* (nonsynonymous nucleotide substitutions) to *Ks* (synonymous nucleotide substitutions) ratios were analyzed via the local package DNASP5.

Expression analysis of *MdCDPK*s based on the transcriptomic data of apple fruits

The RNA-Seq data of both yellow apple 'Blondee' (BLO) and red apple 'Kidd's D-8' (KID) fruits, were downloaded from NCBI with the accession number SRP062637 (http://www.ncbi.nlm.nih.gov/sra). This data set consists of 24 FASTQ files, sequenced for two apple strains (BLO and KID) at the four stages of fruit developments (i.e. early growing-S1, mid growing-S2, late growing-S3, and mature-S4). And each sequencing samples were assigned with three replicates [25]. The analysis of RNA-seq data was carried out following the previous publication [25], except that the HANFU apple genome released in 2019 (https://github.com/moold/Genome-data-of-Hanfu-apple), instead of the apple genome in 2017 [24], was used as the reference genome for reads alignment. Differentially expressed genes (DEGs) were filtered by the $\log_2$FoldChange value more or less than 1 with an adjusted *p*-value ≤ 0.05.

Results

Identification of *CDPK* homologs

A total of 116 candidate CDPK sequences were initially identified from the five species examined, using the CDPK-specific HMMs (Pkinase.hmm and EF-hand_7.hmm) from the Pfam database and searching against the genome data. Among these species, the numbers of CDPK were varied, with 24, 28, 16, 14, and 34 members from apple (*M. domestica*), pear (*P. bretschneideri*), peach (*P. persica*), strawberry (*F. vesca*), and Arabidopsis (*A. thaliana*), respectively. As listed in Table 1, the CDPKs are composed of amino acids ranged from 323 (AT1G76040) ~ 847 (HF10630), with molecular mass from 37.16 ~ 96.22 kDa and pI from 4.49 ~ 9.90.

The identified *CDPK* genes are unevenly distributed on individual genomes (Table 1 and Fig. 1). The 24 apple *CDPK*s (*MdCDPK*s) are distributed among 12 out of 18 chromosomes, including chromosomes no.2 ~ no.7, no.9 ~ no.12, no.14, and no.17, (Fig. 1A). Each of five apple chromosomes (no.3, no.5, no.10, no.12, no.14) contains three *CDPK*s, with the remaining chromosomes having one or two members, respectively. The 28 pear *CDPK*s (*PbCDPK*s) are distributed among 12 out of 17 chromosomes, including chromosomes no.2 ~ no.6, no.9 ~ no.14, and no.17, (Fig. 1B). Among these chromosomes, chromosomes no.12 and no.10 have the most *CDPK*s (five and four, respectively), whereas others have various members ranged from one to three. The 14 strawberry *CDPK*s (*FvCDPK*s) are distributed among 6 out of 7 chromosomes (Fvb2 ~ Fvb7), without *CDPK*s on its chromosome no.1 (Fig. 1C). Chromosome no.6 has the most *CDPK*s (5), in contrast to chromosomes no.4, no.5, and no.7 with the least (1). The 16 peach *CDPK*s (*PpCDPK*s) are distributed throughout all its eight chromosomes (Pp01 ~ 08, Fig. 1D), with the *CDPK* members ranged from one (chromosomes no.2, no.3 and no.6) to four (chromosome no.4). Similarly, the 34 Arabidopsis *CDPK*s (*AtCDPK*s) are found throughout all its five chromosomes (Fig. 1E), with the most *CDPK*s (11) on chromosome no.4 and the least (4) on chromosome no.3. There is a significant uneven distribution of *CDPK*s on the chromosome no.5, due to clustering of six CDPKs on its short arm (Fig. 1E).

Phylogenetic and gene structural analysis of the *CDPK*s

To investigate the phylogenetic relationships and molecular evolutionary history of the sequences in the examined species, following the alignment of 116 CDPK proteins, a phylogenetic analysis was conducted

and a phylogenetic tree was generated using the Maximum Likelihood (ML) method. The phylogenetic tree showed that the 116 CDPKs were clustered into five main subgroups, among which the highest numbers of members were 33 in subgroups I and IV, followed by 31 and 16 in subgroups III and II, respectively (Fig. 2). And subgroup V has the least members (3), all of which are from AtCDPKs. As shown in Fig. 2, the CDPKs from individual species were grouped into different clades rather than a single one. Additionally, their numbers varied within different subgroups. Out of 24 MdCDPKs, nine were located in subgroup I, seven in subgroups IV, and four in subgroups II and III, respectively (Fig. 2). The other *Rosaceae* species also exhibited similar patterns in their CDPK distributions, among which 9, 4, and 4 members from pear, strawberry, and peach, were included into subgroup I, respectively. Accordingly, 5, 2, and 2 in subgroup II; 5, 3, and 6 in subgroup III; 9, 5, and 4 in subgroup (Fig. 2). 34 AtCDPKs were dispersed across subgroups I~V, with the most (13) in subgroup III and the least (3) in subgroups II and V (Fig. 2). Additionally, it appears that AtCDPKs were clustered with each other in the five subgroups, compared to the CDPK clustering across *Rosaceae* species. Moreover, three AtCDPKs were grouped into distinct the subgroup V separated from CDPK homologs from the other species examined (Fig. 2).

Based on sequence alignment, it was found out that all of characteristic domains of CDPK family (i.e. a domain of protein kinase for CDPK activities and four EF-hands for calcium-binding) were presented among 113 out of 116 CDPKs (Fig. 2). Among the remaining 3 CDPKs, AT2G35890 and HF28950 have only two EF-hands (i.e. the 1st and the 2nd ones), whereas Pb001308, without the 1st one, has three EF-hands at the C-terminus, respectively (Fig. 2 and Fig. 3). And there showed a high conservation among the EF-hand domains of the CDPKs identified (Fig. 3).

To characterize their gene structural diversity, the exon-intron organizations of the *CDPK*s were analyzed (Fig. 4A). The number of exons was diverse, with a minimum of one (i.e. *HF00526*, *HF28950*, *Pbr027545*, *Pbr033411*, *Pbr033416*, and *Prupe.3G035400*) and a maximum of 12 (*AT2G17890*, *AT4G04710*, *AT4G36070*, and *AT5G66210*). Generally, the *CDPK* gene structures within each subgroup of the phylogenetic tree, showed a similar pattern, supporting their phylogenetic relationships (Fig. 4A). An exception is that the six *CDPK*s with a single exon are clustered into a distinct clade within subgroup II, which consists of other *CDPK*s with seven exons (8 members), six exons (1member), or two exons (1 member). However, *CDPK*s from a specific clade within a subgroup, apparently have the same numbers of exons (Fig. 4A). In addition, motif analysis by MEME demonstrated that most representatives of the motifs in CDPKs from the same subgroup, showed a conservation in both motif distribution and composition, coordinating with their distribution across various subgroups in the phylogenetic tree (Fig. 4B).

Collinearity analysis of *CDPK*s

To investigate the gene duplication that promotes the evolution of *CDPK* gene family among the species examined, multiple-round analysis of collinearity relationship was carried out between each pair of species. A total of 245 *CDPK* gene-pairs with collinearity relationships were identified, consisting of 36 intraspecies-pairs and 209 interspecies-pairs across each pair of species (Table 2, Fig. 5, Additional file 1: Table S1 and Additional file 2: Figure S1). Among 36 gene-pairs with intraspecies collinearity, 10, 13, 1, and 12 ones were blasted out from apple, pear, peach, and Arabidopsis, respectively (Fig. 5A, 5B and 5C, Table 2 and

Additional file 1: Table S1). And no intraspecies-pair of *CDPK*s was found in strawberry (Fig. 5D). Among 209 gene-pairs with interspecies collinearity, there existed 22, 36, 21, and 25 ones between apple and four other species (i.e. Arabidopsis, pear, strawberry, and peach), respectively (Fig. 5, Table 2 and Additional file 1: Table S1). Both pictorial micro-synteny of 10 *MdCDPK* gene-pairs and 22 *CDPK* gene-pairs between apple and Arabidopsis (Fig. 6), were demonstrated to support their synteny relationships.

Apart from apple and Arabidopsis, *CDPK*s collinearity were also identified between the other species (Additional file 1: Table S1 and Additional file 2: Figure S1). Noticeably, the individual *CDPK* gene-pairs with collinearity relationships were not only distributed within the same phylogenetic subgroups, but also identical in their exon-intron patterns (Fig. 2 and Fig. 4), such as HF05471-Pbr001322, HF20170-Pbr010307, and AT3G10660-AT5G04870 in subgroup I; Pbr033416-Pbr033411, AT1G35670-AT4G09570, and HF39191-Pbr040137 in subgroup II; Pbr024654-Prupe.5G110500, HF17744-Pbr031892, and HF04323-Pbr039714 in subgroup III; HF15429-Pbr021635, HF01706-Pbr036114, and gene25220-Prupe.7G064300 in subgroup IV; AT2G17890-AT4G36070, AT2G17890-AT5G66210, and AT4G36070-AT5G66210 in subgroup V. This result supported the evolutionary relationships between the identified *CDPK*s. To assess the evolutionary rates among these *CDPK* gene-pairs, the *Ka* (nonsynonymous nucleotide substitutions) to *Ks* (synonymous nucleotide substitutions) ratios were calculated (Table 2 and Additional file 1: Table S1). The *Ka*/*Ks* values ranged from 0.022 to 0.751 for the gene-pairs between apple and other species, while from 0.085 to 0.399 for those within apple. It is noticeable that there is no gene-pair with *Ka*/*Ks* values >= 1, inferring that the duplicated *CDPK*s within the species examined have been undergone purifying selection.

Quantitative analysis of *MdCDPK*s expression during apple fruit development

To examined the expression patterns of apple *MdCDPK*s, the expression data set, based on the transcriptomic analysis of two apple strains at their four stages of fruit development [25], was applied to addressing this question. After the quantitative analysis of transcriptomic data, the whole expression levels of 24 *MdCDPK*s were visualized via heatmap plotting (Fig. 7). As showed in Fig. 7, five out of 24 *MdCDPK*s (i.e. *HF03960*, *HF05458*, *HF13700*, *HF20185*, and *HF29516*) presented no any expression amounts at the transcriptional level. In contrast, the remaining 19 *MdCDPK*s were constitutively expressed with various patterns in related to both the apple strains (BLO-yellow fruit skin vs. KID-red fruit skin) and the different stages of fruit development (S1~S4, Fig. 7). The expansion patterns of these *MdCDPK*s were in the trends with three types: (I) higher expression levels, (II) lower expression levels throughout four stages of fruit development, and (III) apparent difference in expression levels at the different stages of fruit development. The expression of six *MdCDPK*s (*HF00526*, *HF05266*, *HF10624*, *HF17744*, *HF36202*, and *HF39191*) were characterized with higher expression levels (i.e. pattern I), while three *MdCDPK*s (*HF04060*, *HF20170*, and *HF28950*) were expressed at lower levels (pattern II). Ten other *MdCDPK*s (*HF01706*, *HF04323*, *HF05471*, *HF06540*, *HF09216*, *HF10630*, *HF13801*, *HF14253*, *HF15429*, and *HF39491*) showed the difference in expression levels at four stages of fruit development (i.e. pattern III), among which the expression levels of *HF01706*, *HF04323*, *HF09216*, and *HF39491*, were in a trend of gradual elevation from the stage S1 to S4, with expression peaks at the stage S4 for both of the apple strains (i.e. BLO and KID). On the contrary, the transcriptional expression in a reverse pattern were found out from another group of *MdCDPK*s (*HF05471*,

HF06540, *HF10630*, *HF13801*, *HF14253*, and *HF15429*), with higher amounts at the S1, compared to those at other stages (Fig. 7).

To further validate if these *MdCDPK*s were differentially expressed with significance among the four stages of fruit development, or the two apple strains, the transcriptional data were screened for the *MdCDPK*s by the criterion: $\log_2$FoldChange > 1 and *p*-value < 0.05 in the present study. As a result, the *MdCDPK*s, differentially expressed with significance, were presented under the twelve types of comparison (Fig. 8). For both the apple strains BLO (Fig. 8A, 8C and 8E) and KID (Fig. 8J, 8K and 8L), the up-regulatory genes (*HF05266* and *HF09216*) and down-regulatory genes (*HF05471* and *HF15429*) were differentially expressed with significance under the comparison S4 vs. S1, S4 vs. S2, and S4 vs. S3. Interestingly, irrespective of their difference in expression fold-changes, the significantly up-regulatory genes (*HF05266* and *HF09216*) and down-regulatory genes (*HF05471* and *HF15429*), were also found out by inter-strain comparison, including KID_S4 vs. BLO_S1, BLO_S2, and BLO_S3 (Fig. 8B, 8D and 8F), or BLO_S4 vs. KID_S1, KID_S2, and KID_S3 (Fig. 8G, 8H and 8I), respectively. In addition, the significantly down-regulatory gene *HF20170* was presented under the comparison S4 vs. S2 for KID (Fig. 8K) or KID_S4 vs. BLO_S1 (Fig. 8B), KID_S4 vs. BLO_S2 (Fig. 8D), while the significantly up-regulatory *HF00526*, *HF04323*, and *HF01706* were under the comparison S4 vs. S2 for BLO (Fig. 8C), S4 vs. S2 for KID (Fig. 8K), and S4 vs. S3 for KID (Fig. 8L), respectively.

Therefore, among the *MdCDPK*s with the expression pattern III at transcriptional level (Fig. 7), five members (*HF01706*, *HF04323*, *HF05471*, *HF09216*, *HF15429*) showed significantly differential expression between the specific stages of fruit development. And the previous described *MdCDPK*s with the expression pattern I (*HF00526*, *HF05266*) or II (*HF20170*) should be sorted to the pattern III, though these *CDPK*s appeared the higher or lower expression amounts throughout the four stages of fruit development due to the limited resolution by the heatmap (Fig. 7).

Discussion

CDPKs play essential roles in modulating a variety of developmental processes, and abiotic stress responses, via mediating Ga^{2+} signatures [6-8]. Although absent in in animals or yeast, CDPKs are widely presented in plants, green algae, and certain protozoa [5]. In the present study, a total of 116 candidate CDPKs were identified in the five species examined, all of which have been validated with the presence of the conserved domains (i.e. kinase and EF-hands, Fig. 2 and Fig. 3). Among these identified *CDPK*s, the maximal number of members was presented in Arabidopsis (34), followed by pear (28), apple (24), and peach (16), while the minimal number in strawberry (14). And the number of *CDPK*s in Arabidopsis, is consistent with those in the previous reports [7]. One of Arabidopsis CDPK (AT2G35890) has been reported to have a truncated C terminus containing two EF-hands, instead of four EF-hands in other AtCDPKs [7]. Analogously, CDPKs with an incomplete of four EF-hands, were showed in one MdCDPK (HF28950 with two EF-hands) and PbCDPK (Pb001308 with three EF-hands), respectively (Fig. 2 and Fig. 3).

Although with a common characteristic of uneven distribution on chromosomes, the *CDPK*s in four species of the *Rosaceae* family do not exhibit a cluster of *CDPK*s on the local region of a chromosome, which could be observed on the short arm of Arabidopsis chromosome no.4 (Fig. 1).

CDPK families in various species consist of a large number of members. For instance, 34 genes encoding *CDPK*s have been revealed from Arabidopsis genome [7], and 40 ones in maize (*Zea mays*) [22], 30 in poplar (*Populus trichocarpa*) [26], 19 in cucumber (*Cucumis sativus*) [27], 27 in barley (*Hordeum vulgare*) [23], 18 in melon (*Cucumis melo*) [28], 98 in upland cotton (*Gossypium hirsutum*) [19]. The generation of *CDPK* multigene family were likely resulted from the whole genome duplication (WGD) [29]. Following the gene duplication, the subsequent evolution was associated with both redundant and distinct functions of CDPK members. According to molecular clock analysis, it was estimated that the diversification of CDPKs in land plants occurred between 268 and 340 MYA (million years ago) [29]. Due to the timing point of divergence between vascular and non-vascular plants (350 ~ 400 MYA), there presented a likelihood that CDPKs in land plants were involved in an adaptation to terrestrial environments [29, 30].

The relationships of 116 CDPKs indicated by the phylogenetic tree were further supported by the similar gene structure and protein-motif patterns within each subgroup. It is noticeable that 34 AtCDPKs were clustered in pairs with each other in all five subgroups of the constructed phylogenetic tree (Fig. 2). In contrast, the clustering CDPK pairs between two different lineages, such as Pbr010295-HF20185, gene05409-Prupe.3G035400, Pbr024654-Prupe.5G110500, were presented among the *Rosaceae* species. The result suggests that AtCDPKs are relatively less close to those in the *Rosaceae* species.

According to the constructed phylogenetic tree (Fig. 2), 24 MdCDPKs were dispersed into the subgroups I (nine MdCDPKs), II (four), III (four), and IV (seven), respectively. In many cases, the *MdCDPK* genes within the same subgroup, do not necessarily present the similar expression patterns at the transcriptional level (Fig. 2, Fig. 7 and Fig. 8). One such case is the significantly up-regulatory *MdCDPK* gene (*HF09216*) and down-regulatory one (*HF15429*) during the four stages of apple fruit development, although both *HF09216* and *HF15429* are from the subgroup IV (Fig. 2, Fig. 7 and Fig. 8). Moreover, another *MdCDPK* gene (*HF29516*) within the subgroup IV, presented no transcriptional expression (Fig. 7). Likewise, within the subgroup III, *HF17744* and *HF04323* showed the expression patterns I and III, respectively, whereas *HF03960* and *HF13700* were in transcriptional silence (Fig. 7). The results indicate that these *CDPK*s have undergone genetic variant events since the evolution of the plant lineage, and potential functional diversification such that single paralogous gene may confer different specificities.

Plants have substantially higher gene duplication rates compared with most other eukaryotes. These plant gene duplicates are mostly derived from tandem, segmental and whole genome duplications. However, the influence of duplication mechanism on *CDPK* gene family in the examined species, has not been thoroughly investigated. To uncover the contribution of gene duplications to the evolution of the *CDPK*s, their collinearity relationships were assessed. Collinearity analysis showed that there presented 36 intraspecies-pairs and 209 interspecies-pairs across each pair of species, including 10, 13, 1, and 12 of intraspecies collinearity within apple, pear, peach, and Arabidopsis, and 22, 36, 21, and 25 of interspecies collinearity between apple with Arabidopsis, pear, strawberry, and peach, respectively (Table 2 and Additional file 1: Table S1). The synteny blocks on the individual chromosome with *CDPK* duplication occurred (Figure 5 and Additional file 1: Figure S1), were in accordance with the large-scale duplication events. Furthermore, according to the analysis on chromosomal locations, the majority of *CDPK*s were unevenly distributed across individual genomes (Fig. 1). Altogether, it is indicated that the WGD duplications might have played

an important role in the *CDPK* gene expansion, leading to structural and functional novelty during evolution of the species lineages.

The *Ka*/*Ks* ratio is considered an indicator for determining the type of selection pressure [31]. In the present research, all of the evaluated *Ka*/*Ks* ratios between *CDPK* gene pairs with collinearity relationships were less than 1 (Table 2 and Additional file 1: Table S1), indicating these genes have undergone purifying selection to different extents since duplication.

Owing to the homologous members of *CDPK* families, it is in concert with no phenotypic effects when one gene in an organism is knocked out. One possible explanation is that the effect of knocking out a gene is compensated by its duplicate copy. With the availability of transcriptome (RNA-seq) as one of alternative methods, the analysis of multigene families, as such *CDPK* gene family, would be promoted in unravelling the functions of a particular *CDPK*. Based on the transcriptomic profiling, Li et al. have reported that three *CDPK*s (*Gb_11259*, *Gb_22778*, *Gb_26648*) were differentially expressed with the significance between the two stages of ovule development in *Ginkgo biloba* [32]. In the present study, the transcriptomic data from RNA-seq analysis on two apple strains at the different stages of fruit development [25], were used for addressing the expression patterns and possible functions of 24 *MdCDPK*s. Apart from five *MdCDPK*s with no expression amounts, 19 *MdCDPK*s, were characterized with three expression patterns according to their heatmap clustering: pattern I and pattern II, with relatively higher and lower expression levels throughout four stages of fruit development, respectively, and pattern III with apparently different expression levels at four stages of fruit development (Fig. 7). Furth significance analysis (log2FoldChange > 1 and p-value < 0.05) on the differential expression of *MdCDPK*s, showed that four *MdCDPK*s, including two up-regulatory expression genes (*HF05266* and *HF09216*) and two down-regulatory genes (*HF05471* and *HF15429*), were differentially expressed with significance under the comparison S4 vs. S1, S4 vs. S2, and S4 vs. S3, respectively (Fig. 8). It is inferred that HF05266 and HF09216, acting as the calcium sensors, may jointly regulate certain downstream targets at the developmental stage S4 in both apple strains (BLO-yellow fruit skin and KID-red fruit skin), whereas HF05471 and HF15429 could be functional at the stages S1, S2, and S3. And other *MdCDPK*s with significantly differential expression, such as the down-regulatory gene *HF20170*, the up-regulatory *HF00526*, *HF04323*, and *HF01706*, were presented only between some of the particular stage comparison (i.e. S4 vs. S2 for KID or KID_S4 vs. BLO_S1, KID_S4 vs. BLO_S2, S4 vs. S2 for BLO, and S4 vs. S3 for KID), indicating their specific roles at the particular stage of apple fruit development. It was reported previously that two *CDPK*s in alfalfa, *MsCK1* and *MsCK2*, were differentially expressed under cold stress, with *MsCK1* down-regulated and *MsCK2* down-regulated, respectively [33], reflecting their specific functions involved in stress responses. With respect to five non-expressed *MdCDPK*s (i.e. *HF03960*, *HF05458*, *HF13700*, *HF20185*, and *HF29516*), there are two assumptions: (1) they are pseudogenes; (2) they are expressed only at specific organs or tissues rather than fruits, or in response to specific stimuli during developmental processes.

Conclusions

A total of 116 *CDPK*s in four *Rosaceae* species (i.e. apple, pear, strawberry, and peach) and Arabidopsis, was characterized at the genome-wide level, including chromosomal distribution, gene structures and conserved

motifs, and phylogenetic and collinearity relationships. These CDPKs were showed to be highly conserved both at their kinase and EF-hand domains. Moreover, the WGD and subsequent purifying selection, might have played an important role in the *CDPK* gene expansion, leading to structural and functional novelty during evolution of the species lineages. Transcriptomic analysis provides an overview for expression patterns of 24 *MdCDPK*s at the four stages of apple fruit development. In many cases, the *MdCDPK* genes within a phylogenetic group could show the different expression patterns at the transcriptional level, suggesting that these *MdCDPK*s have undergone genetic variant events and potential functional diversification such that single paralogous gene may confer different specificities. Furthermore, some of *MdCDPK*s with the fruit developmental stage-specific alteration in expression levels, might be coordinated with their peculiar functions in each case.

Declarations

Acknowledgements

We would like to thank Dr. Junshan Gao at Anhui Agricultural University for his assistance during the genomic data analysis on this research.

Authors' contributions

DHL designed the research project; YD and XXZ, JT, and LFZ carried out the gene family analysis; DHL, YD and XXZ analyzed the transcriptomic data of apple fruit; DHL wrote the paper. All authors read and approved the final manuscript.

Funding

The research, including the design of the study, data collection, analysis and interpretation, and writing manuscript, was supported by the grants from the Natural Science Foundation of Anhui Province, China (No. 1808085MC58). The funders had no other roles during the experiment and preparation of the manuscript.

Availability of data and materials

The gene sequence information mentioned in this study could be found in Table 1, included in the manuscript.

The genome data of *Arabidopsis thaliana*, *Fragaria vesca*, and *Prunus persica*, were download from the database Phytozome (https://phytozome.jgi.doe.gov), those of *Malus domestica* from the online web (https://github.com/moold/Genome-data-of-Hanfu-apple) and *Pyrus bretschneideri* from the Pear Genome Project (http://peargenome.njau.edu.cn), respectively.

Transcriptome information used in this research can be downloaded from the NCBI Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) with the accession number SRP062637.

Other data produced during this work are included in the manuscript and its supplementary files.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

College of Life Science, Anhui Agricultural University, Hefei 230036, China.

References

1. Ludwig AA, Romeis T, Jones JDG. CDPK-mediated signaling pathways: specificity and cross-talk. J Exp Bot. 2004;55:181–8.
2. Sanders D, Pelloux J, Brownlee C, Harper JF. Calcium at the crossroads of signaling. Plant Cell. 2002;14 (Suppl.):S401-17.
3. Dodd AN, Kudla J, Sanders D. The language of calcium signaling. Annu Rev Plant Biol. 2010;61:593–620.
4. DeFalco TA, Bender KW, Snedden WA. Breaking the code: Ca^{2+} sensors in plant signaling. Biochem J. 2010;425:27–40.
5. Hrabak EM, Chan CWM, Gribskov M, Harper JF, et al. The Arabidopsis CDPK-SnRK Superfamily of Protein Kinases. Plant Physiol. 2003;132:666–80.
6. Harmon AC, Gribskov M, Gubrium E, Harper JF. The CDPK superfamily of protein kinases. New Phytol. 2001;151:175–83.
7. Cheng SH, Willmann MR, Chen HC, Sheen J. Calcium signaling through protein kinases. The Arabidopsis calcium-dependent protein kinase gene family. Plant Physiol. 2002;129:469–85.
8. Zhang L, Lu YT. Calmodulin-binding protein kinases in plants. Trends Plant Sci. 2003;8:123–27. DOI: 10.1016/S1360-1385(03)00013-x
9. Martin M, Busconi L. Membrane localization of a rice calcium-dependent protein kinase (CDPK) is mediated by myristoylation and palmitoylation. Plant J. 2000;24:429–35.
10. Evans NH, McAinsh MR, Hetherington AM. Calcium oscillations in higher plants. Curr Opin Plant Biol. 2001;4:415–20.
11. Harper JF. Dissecting calcium oscillators in plant cells. Trends Plant Sci. 2001;6:395–7.
12. Anil VS, Harmon AC, Rao KS. Spatio-temporal accumulation and activity of calcium-dependent kinases during embryogenesis, seed development, and germination in sandalwood. Plant Physiol. 2000;122:1035–44.

13. Yu XC, Zhu SY, Gao GF, Wang XJ, Zhao R, Zou KQ, Wang XF, Zhang XY, Wu FQ, Peng CC, Zhang DP. Expression of a grape calcium-dependent protein kinase ACPK1 in Arabidopsis thaliana promotes plant growth and confers abscisic acid-hypersensitivity in germination, postgermination growth, and stomatal movement. Plant Mol Biol. 2007;64:531–8.
14. Raíces M, Chico JM, Téllez-Iñón MT, Ulloa RM. Molecular characterization of StCDPK1, a calcium-dependent protein kinase from Solanum tuberosum that is induced at the onset of tuber development. Plant Mol Biol. 2001;46:591–601.
15. Yoon GM, Dowd PE, Gilroy S, McCubbin AG. Calcium-dependent protein kinase isoforms in Petunia have distinct functions in pollen tube growth, including regulating polarity. Plant Cell. 2006;18:867–78.
16. Ishida S, Yuasa T, Nakata M, Takahashi Y. A tobacco calcium-dependent protein kinase, CDPK1, regulates the transcription factor REPRESSION OF SHOOT GROWTH in response to gibberellins. Plant Cell. 2008;20:3273–88.
17. Knight H, Knight MR. Abiotic stress signaling pathways: specificity and cross-talk. Trends Plant Sci. 2001;6:262–7.
18. Chico JM, Raíces M, Téllez-Iñón MT, Ulloa RM. A calcium-dependent protein kinase is systemically induced upon wounding in tomato plants. Plant Physiol. 2002;128:256–70.
19. Gao W, Xu FC, Guo DD, Zhao JR, Liu J, Guo YW, Singh PK, Ma XN, Long L, Botella JR, Song CP. Calcium-dependent protein kinases in cotton: insights into early plant responses to salt stress. BMC Plant Biol. 2018;18:15.
20. Romeis T, Ludwig AA, Martin R, Jones JDG. Calcium-dependent protein kinases play an essential role in a plant defence response. EMBO J. 2001;20:5556–67.
21. Saijo Y, Hata S, Kyozuka J, Shimanoto K, Izui K. Over-expression of a single Ca^{2+}-dependent protein kinase confers both cold and salt/drought tolerance on rice plants. Plant J. 2000;23:319–27.
22. 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 in maize. BMC Genomics. 2013;14:433.
23. Fedorowicz-Strońska O, Koczyk G, Kaczmarek M, Krajewski P, Sadowsk J. Genome-wide identification, characterization and expression profiles of calcium-dependent protein kinase genes in barley (*Hordeum vulgare* L.). J Appl Genet. 2017;58:11–22.
24. Daccord N, Celton JM, Linsmith G, Becker C, et al. High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development. Nat Genet. 2017;49:1099–108.
25. El-Sharkawy I, Liang D, Xu K. Transcriptome analysis of an apple (*Malus* × *domestica*) yellow fruit somatic mutation identifies a gene network module highly associated with anthocyanin and epigenetic regulation. J Exp Bot. 2015;66:7359–76.
26. Zuo R, Hu R, Chai G, Xu M, Qi G, Kong Y, Zhou G. Genome-wide identification, classification, and expression analysis of CDPK and its closely related gene families in poplar (*Populus trichocarpa*). Mol Biol Rep. 2013;40:2645–62.
27. 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 and Genomics,

2015;290:1403–14.

28. Zhang H, Wei C, Yang X, Chen H, Yang Y, Mo Y, Li H, Zhang Y, Ma J, Yang J, Zhang X. Genome-wide identification and expression analysis of calcium-dependent protein kinase and its related kinase gene families in melon (*Cucumis melo* L.). PLoS ONE. 2017;12:e0176352.
29. Valmonte GR, Arthur K, Higgins CM, MacDiarmid RM. Calcium-dependent protein kinases in plants: evolution, expression and function. Plant Cell Physiol. 2014;55:551–69.
30. Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D. A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol. 2004;21:809–18.
31. Wagner A. Selection and gene duplication: a view from the genome. Genome Biol. 2002;3:reviews1012.1.
32. Li DH, Wu D, Li SZ, Guo N, Gao JS, Sun X, Cai YP. Transcriptomic profiling identifies differentially expressed genes associated with programmed cell death of nucellar cells in *Ginkgo biloba* L. BMC Plant Biol. 2019;19: 91.
33. Monroy AF, Dhindsa RS. Low-temperature signal transduction: induction of cold acclimation- specific genes of alfalfa by calcium at 25 degrees C. Plant Cell. 1995;7:321–31.
34. Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40:e49.
35. Li DH, Wu D, Li SZ, Dai Y, Cao YP. Evolutionary and functional analysis of the plant-specific NADPH oxidase gene family in *Brassica rapa* L. Roy Soc Open Sci. 2019;6:181727.

Tables

Table 1 116 *CDPK* genes identified in the five species.

Species	GeneID	Chr.	Start	End	A.A._length	Mol._Wt.	pI
Apple	HF00526	Chr09	14717392	14718888	498	55509.25	5.01
	HF01706	Chr14	27337518	27340944	569	64834.37	7.44
	HF03960	Chr03	21504700	21508486	589	65585.14	6.07
	HF04060	Chr03	23624970	23638053	623	69689.59	6.56
	HF04323	Chr03	27209001	27213712	544	61047.43	6.87
	HF05266	Chr12	8231156	8234728	594	67191.96	6.9
	HF05458	Chr12	4856269	4863443	636	70998.67	6.22
	HF05471	Chr12	4677427	4680675	657	73313.49	6.17
	HF06540	Chr02	31814438	31830681	671	75356.91	6.48
	HF09216	Chr06	31360009	31363493	570	64338.99	7.03
	HF10624	Chr05	27714106	27717172	571	64077.56	6.63
	HF10630	Chr05	27734894	27747686	847	96220.08	6.11
	HF13700	Chr11	23569338	23571980	533	59745.47	5.79
	HF13801	Chr11	26393433	26396848	534	59844.21	6.57
	HF14253	Chr07	5890600	5895717	577	64203.12	5.19
	HF15429	Chr04	22910643	22915419	527	59440.69	6.29
	HF17744	Chr10	42130315	42133600	547	61187.61	6.68
	HF20170	Chr14	4323883	4326917	647	71860.72	6.27
	HF20185	Chr14	4465648	4471454	611	67916.23	6.1
	HF28950	Chr17	13770005	13770997	330	37365.92	5.47
	HF29516	Chr17	6974338	6976919	541	61651.21	5.26
	HF36202	Chr10	23807977	23811730	573	64410.54	6.13
	HF39191	Chr10	5715948	5719528	500	56155.17	5.62
	HF39491	Chr05	7971543	7975169	495	55634.4	5.29
Pear	Pbr001308	Chr12	18052931	18058996	548	61017.27	6.47
	Pbr001322	Chr12	18207189	18210436	657	73639.95	6.11
	Pbr006943	Chr11	9792223	9796137	533	59771.35	6.25
	Pbr010295	Chr14	1609323	1614889	604	67234.36	6.1
	Pbr010307	Chr14	1740855	1743888	647	71622.31	6.19

	Pbr011500	Chr6	1375205	1379132	549	61973.21	6.99
	Pbr011659	Chr6	18021952	18025217	533	59623.01	6.25
	Pbr017213	Chr3	8945299	8949047	533	59619.95	6.25
	Pbr018253	Chr10	24089960	24093533	521	58594.09	5.73
	Pbr018323	Chr13	6403652	6405694	338	38528.49	4.51
	Pbr021635	Chr4	3825488	3831583	527	59440.69	6.29
	Pbr023408	Chr12	14943714	14947912	531	59826.19	5.91
	Pbr023960	Chr5	15139232	15143827	571	63918.32	6.35
	Pbr024654	Chr6	9640358	9645265	525	58944.11	6.52
	Pbr027545	Chr9	10270428	10272683	498	55628.3	4.86
	Pbr028710	Chr13	1306362	1315376	580	65229.99	4.49
	Pbr028878	Chr12	2768254	2772775	527	59342.63	6.41
	Pbr028879	Chr12	2852127	2856648	527	59342.63	6.41
	Pbr029596	Chr11	15933676	15936752	534	59754.51	5.92
	Pbr031892	Chr10	5306484	5309784	546	61047.36	6.52
	Pbr032128	Chr10	13533994	13539494	570	63740.96	6.18
	Pbr033297	Chr5	14432235	14436411	571	63918.32	6.35
	Pbr033365	Chr2	10536286	10545796	739	82644.04	4.73
	Pbr033411	Chr17	13856646	13858154	502	55933.64	4.91
	Pbr033416	Chr17	13738242	13739750	502	55933.64	4.91
	Pbr036114	Chr14	17344464	17348428	548	62308.56	7.2
	Pbr039714	Chr3	6216828	6220557	544	61061.47	6.8
	Pbr040137	Chr10	23595121	23598690	521	58594.09	5.73
Peach	Prupe.1G190700	Pp01	17596846	17600318	543	61699.47	5.77
	Prupe.1G360100	Pp01	33227048	33230862	534	60384.65	6.34
	Prupe.2G075100	Pp02	11450539	11458447	567	63193.96	5.19
	Prupe.3G035400	Pp03	2573217	2575613	502	56527.43	4.97
	Prupe.4G021300	Pp04	1001928	1006361	552	61762.43	6.71
	Prupe.4G213800	Pp04	13353107	13358433	545	61168.54	6.92
	Prupe.4G250200	Pp04	16736364	16741425	533	59914.49	6.18

	Prupe.4G263600	Pp04	19357207	19360505	534	59759.52	5.8
	Prupe.5G110500	Pp05	11501794	11507019	526	59008.23	6.47
	Prupe.5G184800	Pp05	15348366	15352834	548	62195.44	7.37
	Prupe.6G278500	Pp06	26117102	26122709	527	59443.8	6.29
	Prupe.7G064300	Pp07	10277008	10282004	531	60208.02	6.8
	Prupe.7G089600	Pp07	12322567	12329354	605	67284.36	5.15
	Prupe.7G090800	Pp07	12422089	12426248	639	71300.01	6.25
	Prupe.8G057600	Pp08	7106577	7111590	497	55731.47	4.93
	Prupe.8G180800	Pp08	18124933	18129600	573	64304.92	6.39
Strawberry	gene01742	Fvb4	915918	924111	536	61005.67	5.42
	gene05409	Fvb6	27595413	27597261	466	52474.8	4.79
	gene08576	Fvb2	27160916	27165535	533	60124.48	6
	gene09567	Fvb5	9036470	9040723	547	62205.47	6.91
	gene13451	Fvb6	6760540	6766507	527	59234.57	6.47
	gene14687	Fvb3	21890269	21896822	538	59957.38	6.22
	gene15357	Fvb3	21000515	21005183	531	59297.16	5.73
	gene17341	Fvb2	14313210	14318026	568	63588.15	6.08
	gene18135	Fvb6	17076865	17083149	595	65930.5	4.91
	gene18254	Fvb6	17276867	17281015	635	70545.92	5.52
	gene19615	Fvb3	1053149	1056363	543	61074.39	6.57
	gene23668	Fvb7	5239400	5246458	562	62545.37	5.34
	gene25220	Fvb6	18554274	18558187	529	59602.96	6.57
	gene27440	Fvb2	2846299	2849227	490	55346.41	5.5
Arabidopsis	AT1G18890	Chr1	6522764	6525962	545	61459.77	6.6
	AT1G35670	Chr1	13205381	13208252	495	55915.97	5.08
	AT1G50700	Chr1	18781914	18784582	521	58605.82	6.38
	AT1G61950	Chr1	22899417	22901946	551	62948.22	7.03
	AT1G74740	Chr1	28079946	28082644	541	61404.47	6.71
	AT1G76040	Chr1	28538830	28540637	323	37161.68	5.11
	AT2G17290	Chr2	7516415	7519633	544	61111.47	5.39

	AT2G17890	Chr2	7769885	7772627	571	64753.95	9.9
	AT2G31500	Chr2	13413764	13416536	582	66243.93	7.4
	AT2G35890	Chr2	15067175	15069136	520	58851.84	5.66
	AT2G38910	Chr2	16245214	16247483	583	64720.58	5.25
	AT2G41860	Chr2	17467344	17469786	530	60054.13	7.21
	AT3G10660	Chr3	3331398	3334268	646	72254.06	5.1
	AT3G20410	Chr3	7116201	7119121	541	60362.56	6.06
	AT3G51850	Chr3	19232467	19235889	528	59375.89	6.65
	AT3G57530	Chr3	21296554	21299591	538	60935.45	6.36
	AT4G04695	Chr4	2381634	2383996	484	54696.71	6.32
	AT4G04700	Chr4	2385276	2387986	485	54898.53	4.79
	AT4G04710	Chr4	2389598	2392887	575	64674.87	5.22
	AT4G04720	Chr4	2394458	2397759	531	59894.46	6.64
	AT4G04740	Chr4	2403609	2408737	520	58654.03	6.52
	AT4G09570	Chr4	6049517	6052335	501	56416.37	5.02
	AT4G21940	Chr4	11640807	11643641	554	62575.4	6.09
	AT4G23650	Chr4	12324758	12327459	529	59336.29	6.32
	AT4G35310	Chr4	16801987	16804995	556	62127.31	4.97
	AT4G36070	Chr4	17056907	17059595	534	60213.52	8.94
	AT4G38230	Chr4	17928677	17931182	340	38204.5	4.65
	AT5G04870	Chr5	1416783	1420338	610	68253.67	5.18
	AT5G12180	Chr5	3937024	3939596	528	58484.4	5.96
	AT5G12480	Chr5	4047515	4050533	535	60309.67	6.13
	AT5G19360	Chr5	6521716	6523780	523	58174.08	5.77
	AT5G19450	Chr5	6558426	6561534	533	59940.53	6.3
	AT5G23580	Chr5	7950202	7952532	490	55379.27	5.08
	AT5G66210	Chr5	26456292	26459624	523	58972.11	8.65

Table 2 *Ka*/*Ks* analysis for the *CDPK* gene pairs with collinearity relationship within apple or between apple and Arabidopsis.

Gene Pairs	*Ks*	*Ka*	*Ka/Ks*	Purifying selection
HF00526-HF28950	0.27205	0.06238	0.229	Yes
HF03960-HF13700	0.16370	0.03628	0.222	Yes
HF04060-HF13801	0.31354	0.03178	0.101	Yes
HF04323-HF17744	1.98321	0.16921	0.085	Yes
HF05458-HF20185	0.13035	0.02240	0.172	Yes
HF05471-HF20170	2.76653	0.30587	0.111	Yes
HF06540-HF14253	0.17269	0.02159	0.125	Yes
HF09216-HF01706	0.16524	0.03024	0.183	Yes
HF10630-HF36202	0.30008	0.11987	0.399	Yes
HF39491-HF39191	0.16721	0.02202	0.132	Yes
AT1G18890-HF01706	1.76044	0.11626	0.066	Yes
AT1G18890-HF09216	1.81065	0.39689	0.219	Yes
AT1G61950-HF17744	1.84570	0.21937	0.119	Yes
AT1G74740-HF01706	2.88293	0.12300	0.043	Yes
AT1G74740-HF09216	1.71041	0.13587	0.079	Yes
AT2G17290-HF10630	2.74471	0.17422	0.063	Yes
AT2G17290-HF36202	2.05040	0.09918	0.048	Yes
AT2G35890-HF05471	1.75350	0.10159	0.058	Yes
AT2G35890-HF20170	2.41659	0.29969	0.124	Yes
AT2G41860-HF05266	2.20495	0.16753	0.076	Yes
AT3G20410-HF04323	3.71499	0.19827	0.053	Yes
AT3G51850-HF15429	1.75018	0.08026	0.046	Yes
AT3G57530-HF05266	2.20618	0.14990	0.068	Yes
AT4G04710-HF17744	2.73536	0.45878	0.168	Yes
AT4G21940-HF17744	3.54523	0.20960	0.059	Yes
AT4G35310-HF10630	2.24867	0.16539	0.074	Yes
AT4G35310-HF36202	1.88606	0.11305	0.060	Yes

AT4G38230-HF10624	2.19955	0.11275	0.051	Yes
AT4G38230-HF36202	1.80224	0.11446	0.064	Yes
AT5G12480-HF04060	2.15020	0.12973	0.060	Yes
AT5G12480-HF13801	2.39258	0.12622	0.053	Yes
AT5G19360-HF13700	3.07646	0.13074	0.042	Yes

Figures

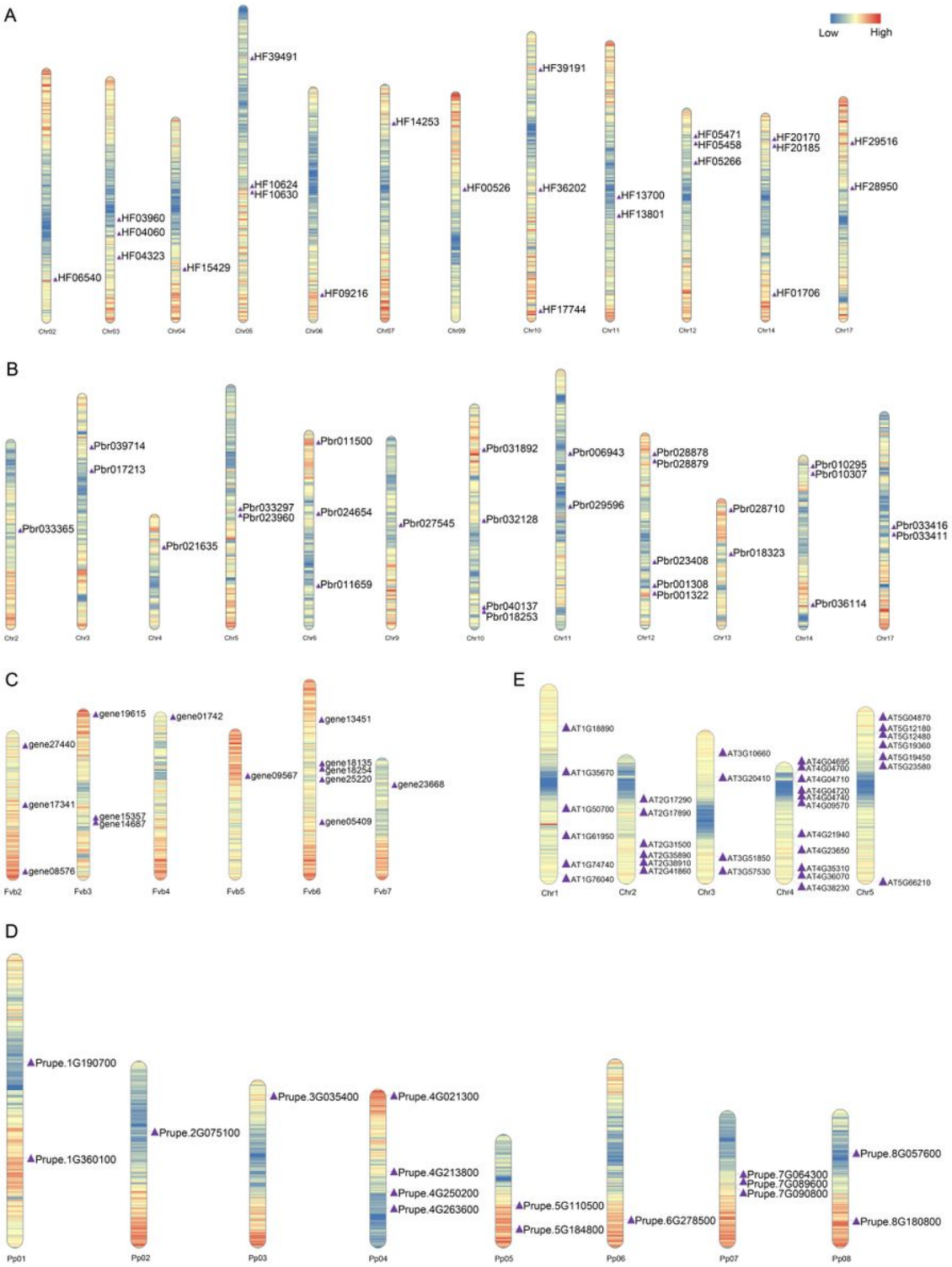

Figure 1

Chromosomal locations of CDPK genes according to the individual physical position. A. MdCDPKs; B. PbCDPKs; C. FvCDPKs; D. PpCDPKs; E. AtCDPKs. The color bar from blue to red indicates the gene density of respective chromosomes.

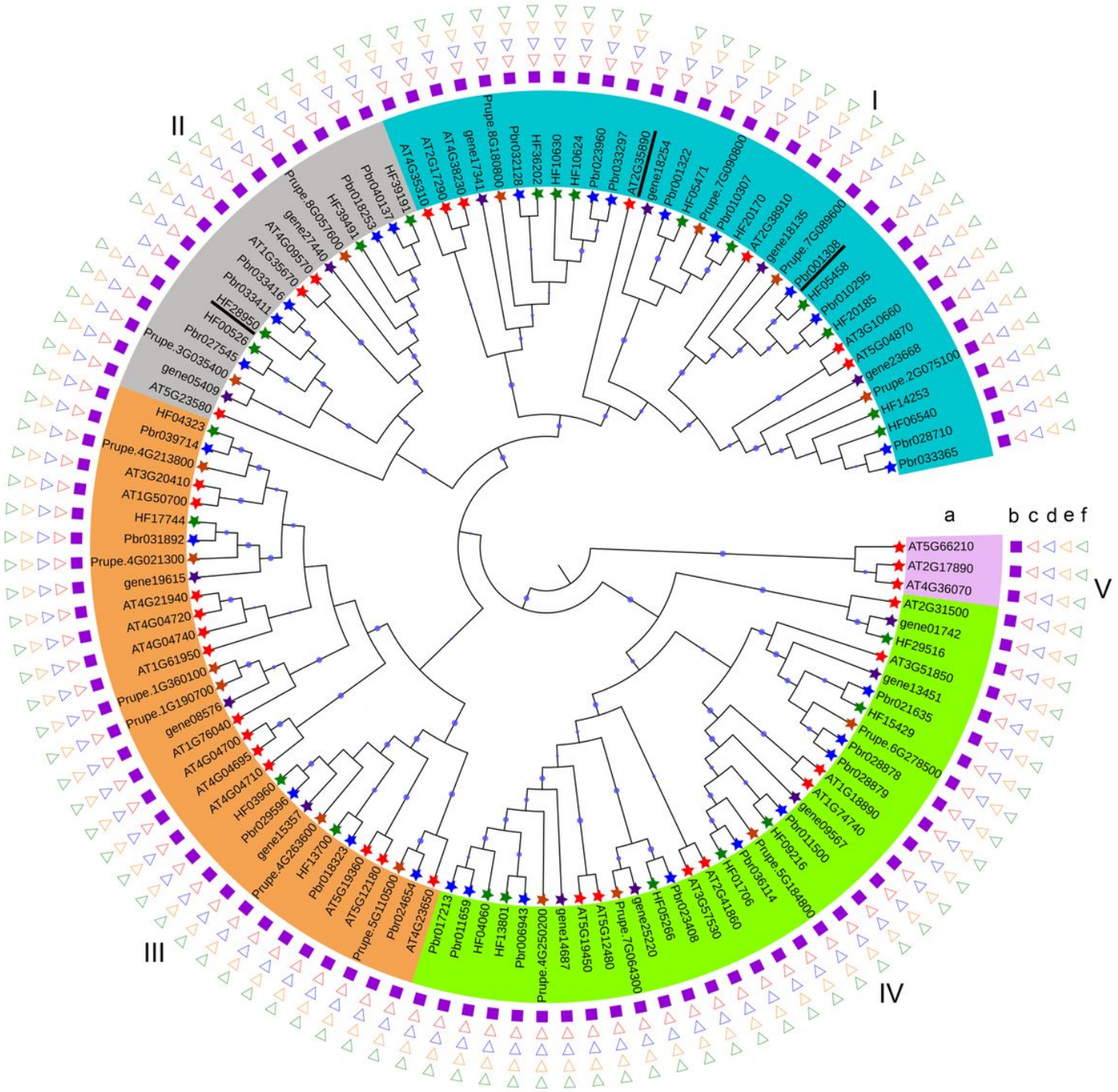

Figure 2

Phylogenetic tree constructed based on the amino acid sequences of 116 CDPKs. Five subgroups are indicated with I–V, respectively. Square in purple refer to the kinase domains. Triangles in red, blue, yellow, and green indicate the domains of four EF hands, respectively. Circle a, marks the CDPK IDs. Circles b, c, d, e, and f are composed of the above-mentioned domains, respectively. Circles at the individual nodes represent bootstrap support. Three CDPKs with the incomplete EF hands are underlined.

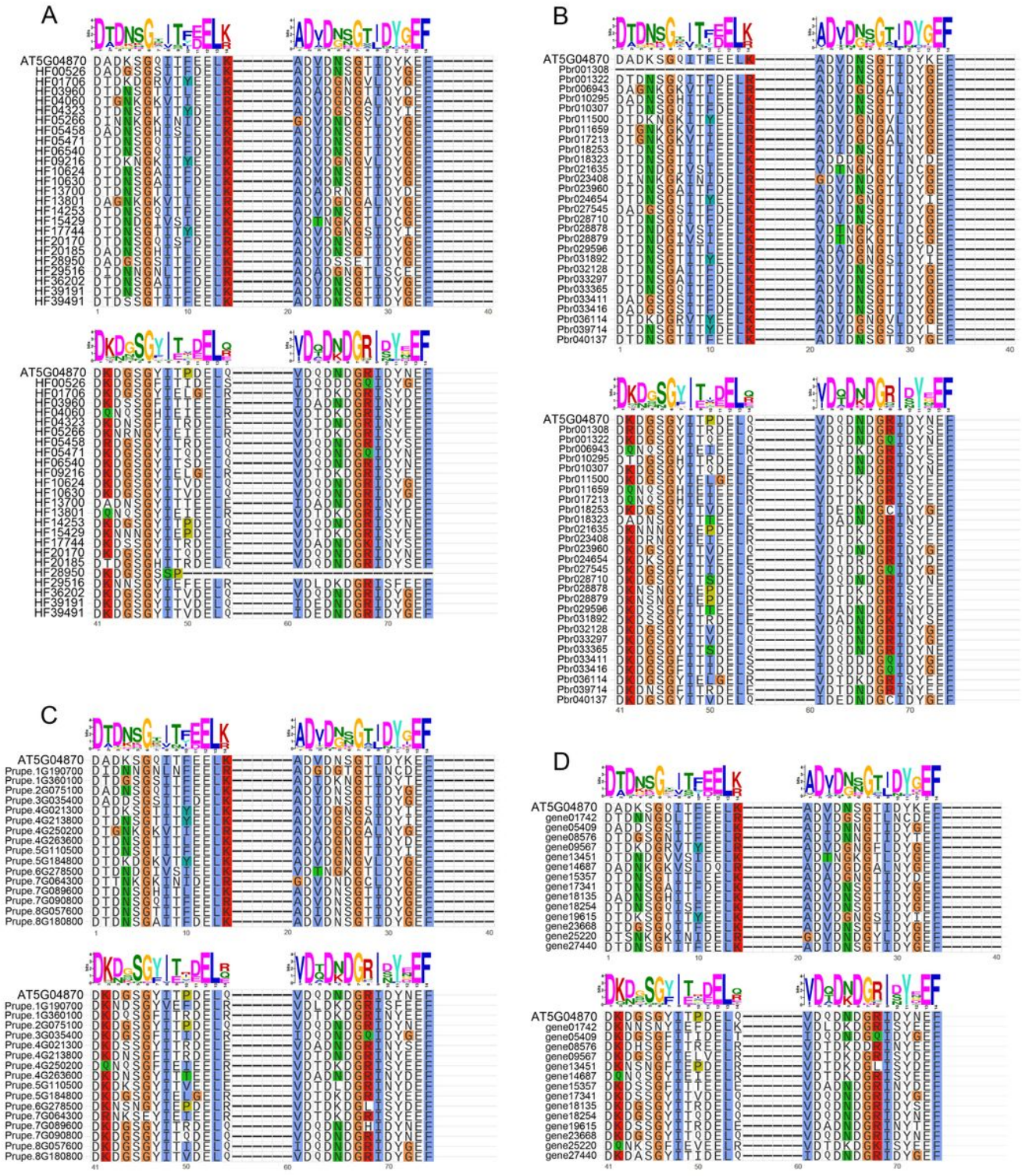

Figure 3

Sequence logos of the conserved domains of four EF-hands. A. MdCDPKs; B. PbCDPKs; C. PpCDPKs; D. FvCDPKs. The alignments were performed using the amino acid sequences of four EF-hands from a AtCDPK, AT5G04870, as the reference.

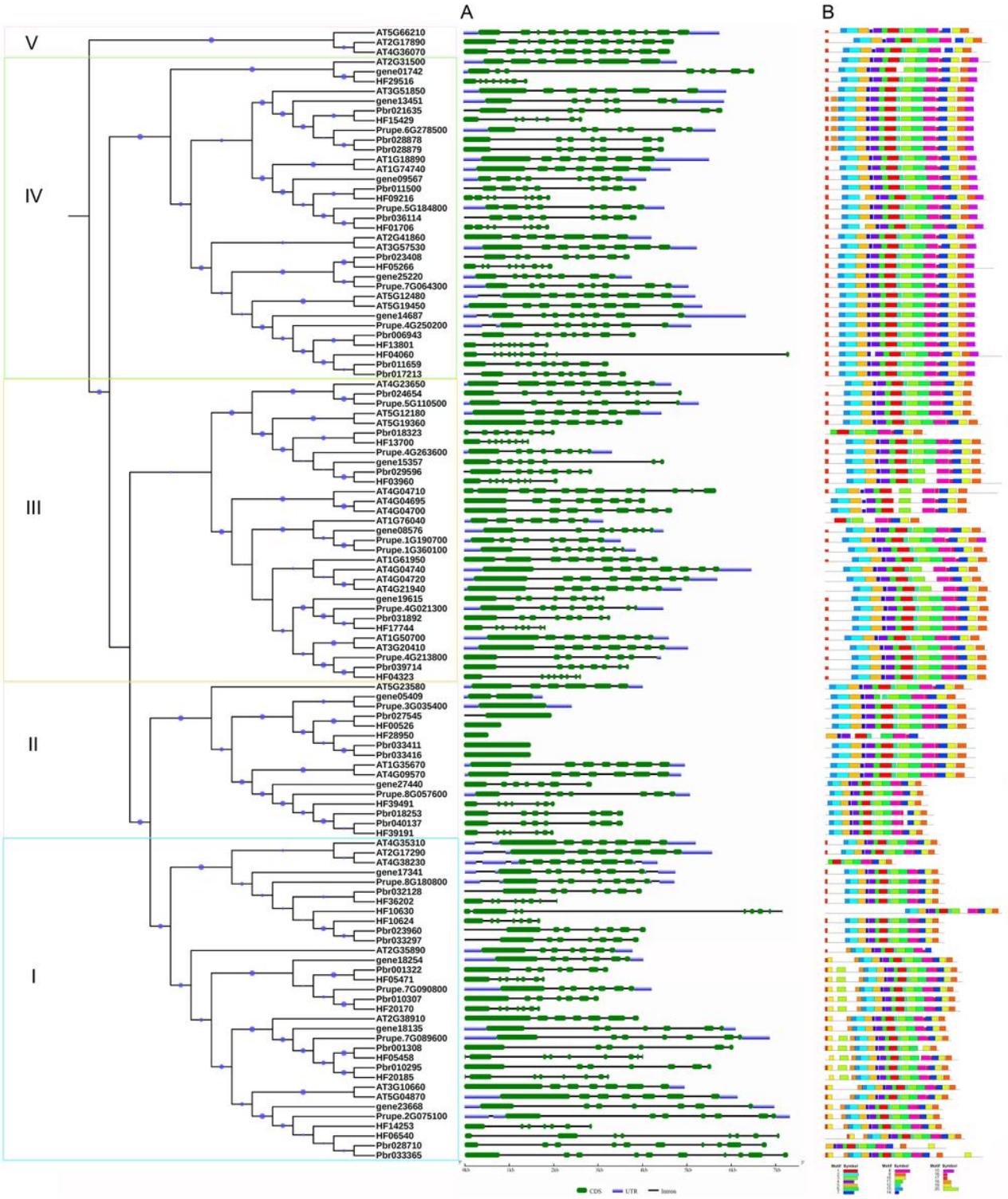

Figure 4

Gene structures (A) and protein motifs (B) of CDPKs. The markers I–V indicate the five corresponding subgroups. Blue boxes: 5’ or 3’ untranslated region (UTR); green boxes: exons; black lines: introns. Protein motifs are analyzed by MEME program, with boxes in colors indicating 20 motifs.

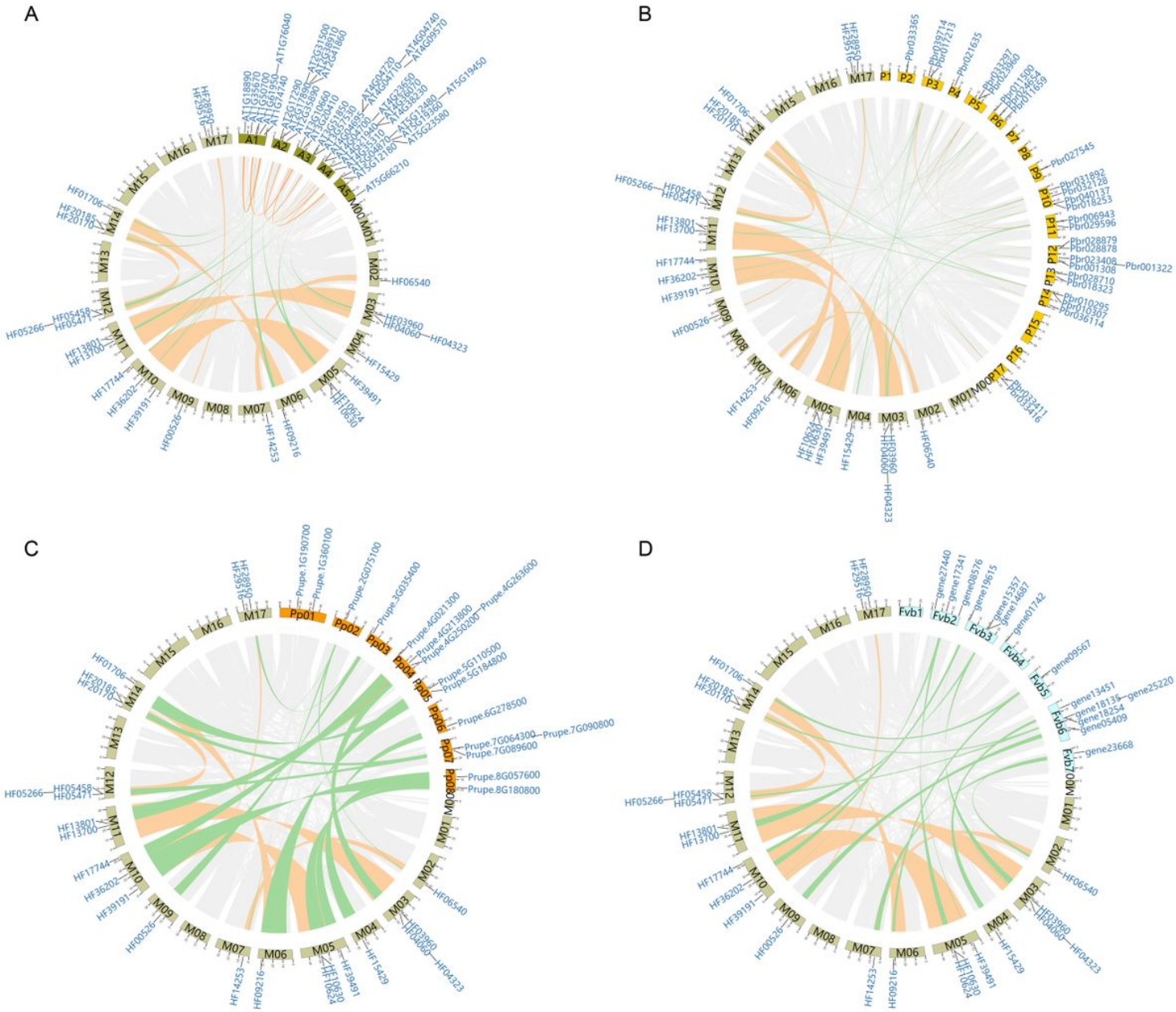

Figure 5

Collinearity relationships between MdCDPKs and AtCDPKs (A), PbCDPKs (B), PpCDPKs (C), and FvCDPKs (D). Chromosomes of Arabidopsis, apple, pear, peach and strawberry, are marked with A1–A5, M00–M17, P1–P17, Pp01–Pp08, and Fvb1–Fvb7, respectively. Orange linkages indicate the intraspecies synteny blocks, while green ones refer to the interspecies synteny blocks, which contain the CDPK gene-pairs with collinearity. The gene IDs labeled on chromosomes are CDPKs identified in the five species.

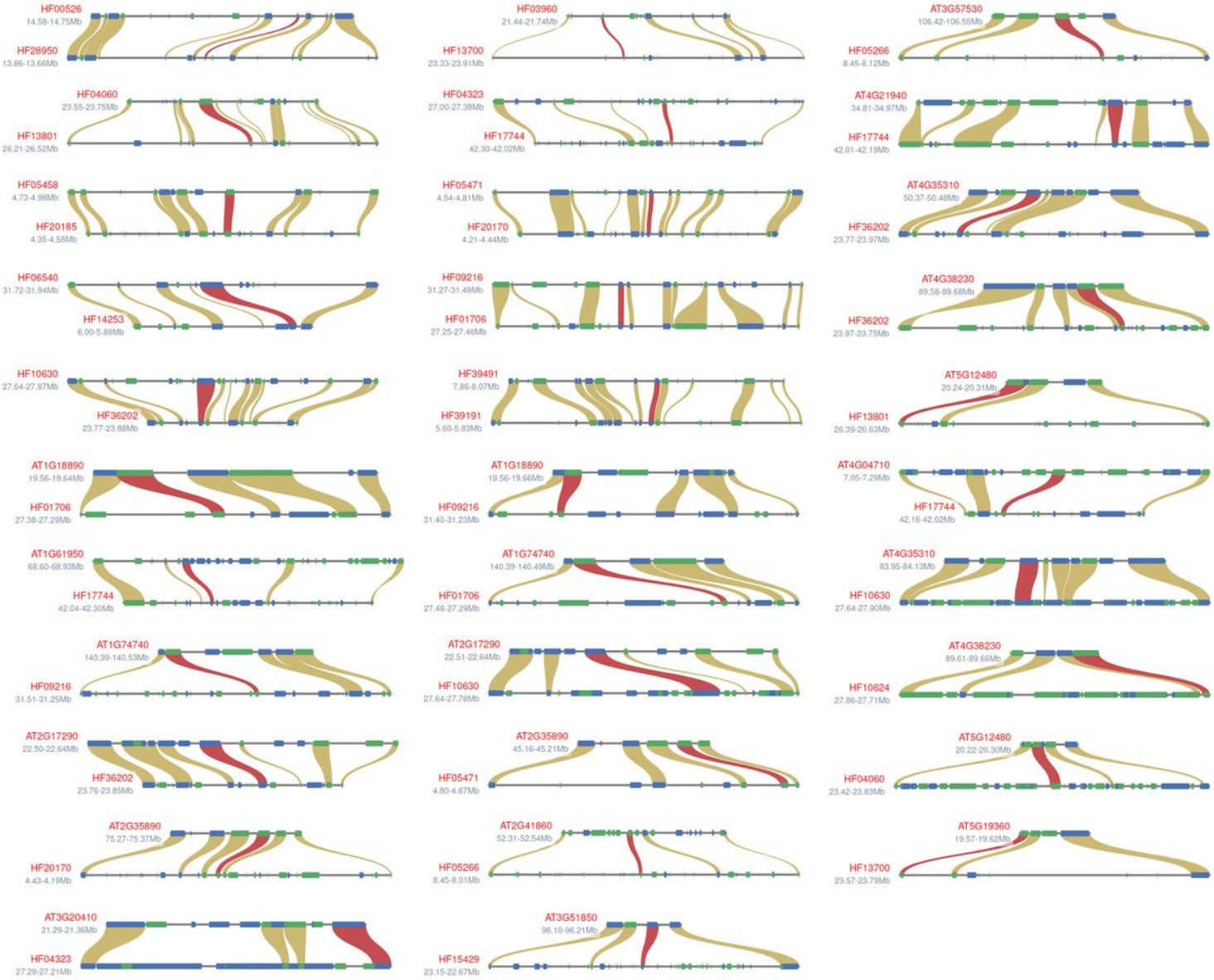

Figure 6

Micro-synteny of CDPKs within apple, or between apple and Arabidopsis. A region less than 0.1 Mb flanking at both sides of CDPKs is illustrated. The grey horizontal lines represent the chromosomal segments, while blue and green boxes represent the genes with cis and trans direction, respectively. CDPK gene pairs are linked by the red curves, whereas other gene pairs are connected by the golden ones. IDs of the CDPK gene pairs are listed at the upstream of chromosomal segments.

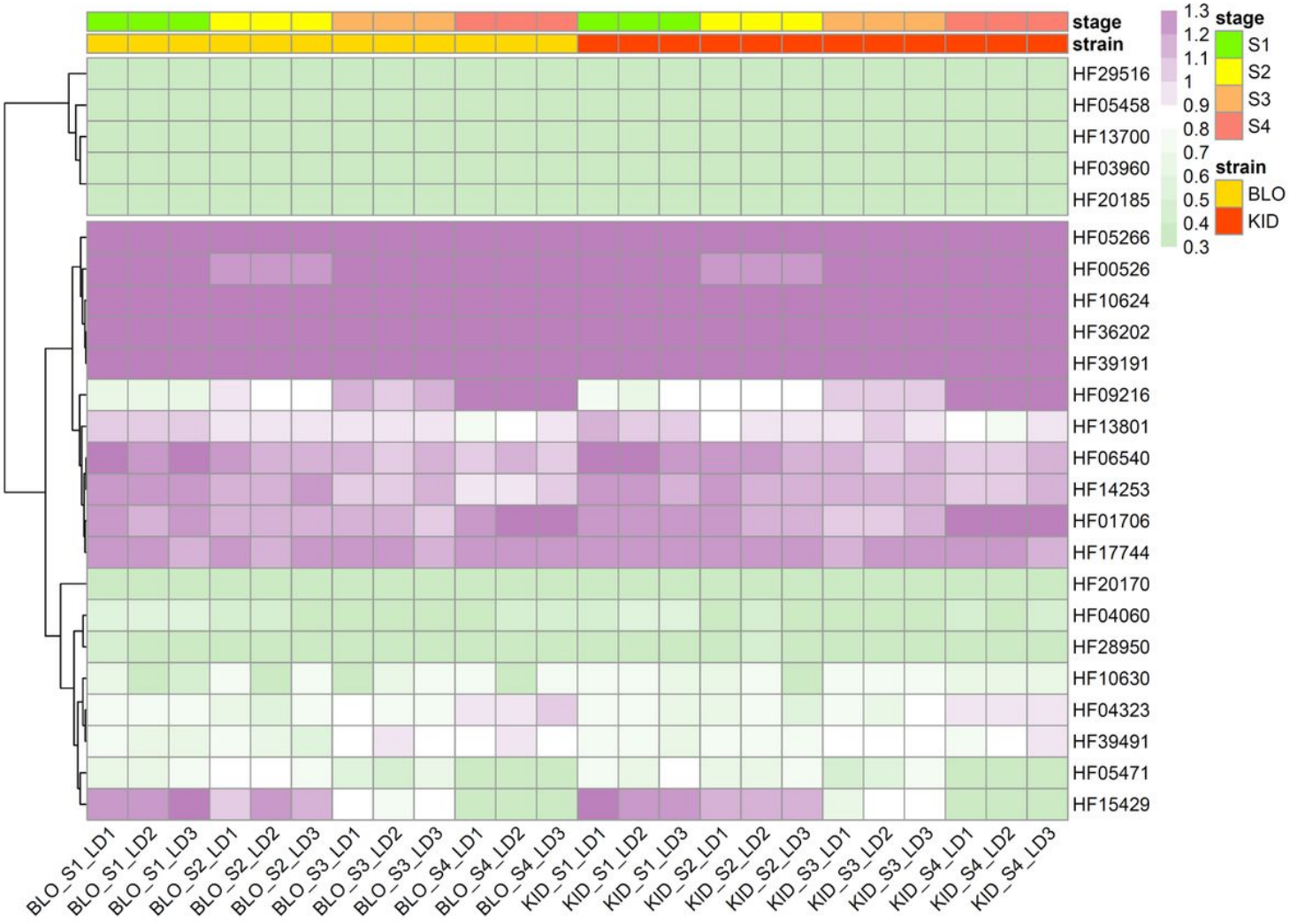

Figure 7

Heatmaps of expression patterns of 24 MdCDPKs for two apple strains (BLO and KID) at the four stages (S1–S4) of apple fruit development. LD1, LD2, and LD3 are triplicates of the RNA-seq samples.

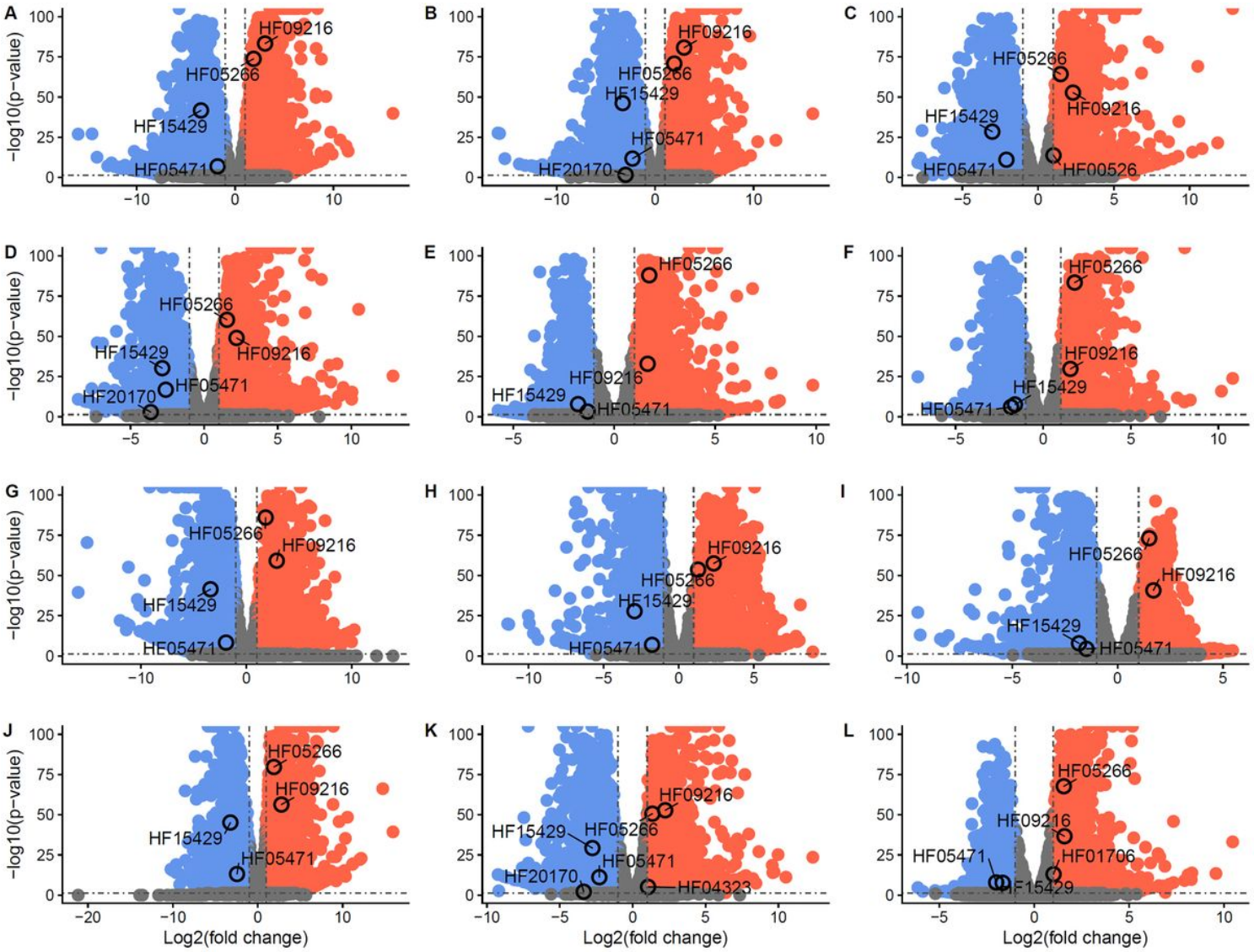

Figure 8

The expression profiles to identify the differentially expressed genes (DEGs) in the RNA-seq. Red and blue dots refer to the significant DEGs with p-value ≤ 0.05 and log2(foldchange) > 1, while grey ones indicate those without significance, respectively. The differentially expressed CDPK genes with significance, are marked by black circles under the comparison BLO_S4 vs. BLO_S1 (A), BLO_S4 vs. BLO_S2 (C), BLO_S4 vs. BLO_S3 (E), KID_S4 vs. KID_S1 (J), KID_S4 vs. KID_S2 (K), KID_S4 vs. KID_S3 (L), KID_S4 vs. BLO_S1 (B), KID_S4 vs. BLO_S2 (D), KID_S4 vs. BLO_S3 (F), BLO_S4 vs. KID_S1 (G), BLO_S4 vs. KID_S2 (H), and BLO_S4 vs. KID_S3 (I), respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1TableS1.xlsx
- Additionalfile2FigureS1.jpg