Identification, evolution and functional inference on the cold-shock domain protein family in Pak-choi (Brassica rapa ssp. chinensis) and Chinese cabbage (Brassica rapa ssp. pekinensis)

Feiyi Huang*, Jin Wang*, Jun Tang and Xilin Hou

State Key Laboratory of Crop Genetics & Germplasm Enhancement, Key Laboratory of Biology and Genetic Improvement of Horticultural Crops (East China), Ministry of Agriculture and Rural Affairs of the P.R.China, Engineering Research Center of Germplasm Enhancement and Utilization of Horticultural Crops, Ministry of Education of the P.R.China, Nanjing, 210095, People’s Republic of China

ABSTRACT
The cold shock domain proteins (CSPs) play important roles in plant developmental processes and stress responses. Multiple sequence alignments found that these CSPs all have a cold shock domain fragment. Phylogenetic tree analysis found that most BcCSPs were more closely related to BrCSPs than other crops. Furthermore, we observed the conserved intron/exon structural patterns of these CSP genes in Pak-choi, Chinese cabbage and Arabidopsis, all motif were observed to be conserved among several boxes subgroups. Comparative analysis of expression patterns of CSP genes in Pak-choi, Chinese cabbage and Arabidopsis suggest that the CSP genes play various roles in plants through qRT-PCR analysis. Further, the CSP expression pattern in different tissues (roots, stems and leaves) of Chinese cabbage was also studied. This study showed that these gene family members might play roles in abiotic stresses responses, and might benefit from their functional characterization and utilization in the resistance engineering of Pak-choi and Chinese cabbage.

Introduction
Low temperatures adversely affect plant growth and development (Winfield et al. 2010). During the process of long-term adaptive evolution, some temperate plant species have evolved a series of physiological and molecular mechanisms to acclimate to low temperatures (Song et al. 2016). Cold shock domain proteins (CSPs) are among the most conserved nucleic acid-binding proteins consisting of a small gene family in plants. CSPs participate in various cellular functions, which are mediated by their characteristic ability to bind nucleic acids (Thompson 2008; Chaikam and Karlson 2010; Barnes et al. 2016). CSPs are essential for organisms to grow and survive in cold environments, CSPs might play important roles in abiotic stress, as low temperature stress, salt stress and so on (Zhang et al. 2017).

We conducted this experiment the identification and molecular cloning of CSP gene family members in Pak-choi, Chinese cabbage and Arabidopsis. Among the four AtCSPs, AtCSP2 (AtGRP2/CSDP2; At4g38680) is characterized for its function in affecting flowering time and reproductive tissue development, including seed development (Fusaro et al. 2007). Plant stress responses are generally controlled by a network of specialized genes through intricate regulation by specific transcription factors (Nuruzzaman et al. 2010). The CSP gene family, an important member of the stress-related transcription factor families, is involved in the regulation of plant developmental processes and biotic and abiotic stress. CSPs have been recognized in representative dicots, monocots and mosses (Karlson 2008; Karlson et al. 2009; Chaikam and Karlson 2010). The Arabidopsis Columbia (Col-0) genome has four genes that encode a protein with an N-terminal CSD domain, multiple CCHC zinc fingers and glycine-rich regions (AtCSP1-AtCSP4). CSP1 and CSP3 are similar in domain organization, as are CSP2 and CSP4. CSPs are also implicated in abiotic stress responses as AtCSP1-AtCSP3 transcripts were enhanced by cold treatment (Karlson et al. 2002), and over expression of AtCSP3 but not AtCSP1 or AtCSP2 improved freezing tolerance (Kim et al. 2009). Arabidopsis CSP1 binds RNA in vitro, leading to the proposal that it functions as an RNA chaperone, however, it is not known if CSPs selectively bind individual mRNAs or facilitate specific post-transcriptional processes (Kim et al. 2009; Li 2014). In this study, CSP genes were obtained from Pak-choi, Chinese cabbage and Arabidopsis. Based on a full-length cDNA sequence we obtained first, we conducted family and evolution analysis, CSPs expression level were then exposed to temperature shocks and at different abiotic stress. We believe the results provide clues to further decipher CSP functions in Pak-choi, Chinese cabbage and Arabidopsis.

Materials and methods
Plant material, growth conditions and abiotic stress treatments
Pak-choi (Brassica rapa ssp. chinensis cv. Suzhouqing) and Chinese cabbage (B. rapa ssp. pekinensis), obtained from the Non-heading Chinese cabbage project team in Nanjing Agricultural University. Healthy seeds were harvested from the same batch, which were grown in an artificial climate
chamber at 24°C under continuous cool white fluorescent illumination at an intensity of 100 μmol m⁻² s⁻¹ (Gilmour et al. 1998). Then those seeds were placed in a growth chamber (16 h 22°C, light; 8 h 20°C, dark daily). Stress treatments were performed using the seedlings at their five-leaf stage and all treatments (ABA, Cold, Salt, and Osmotic) were processed under hydroponic conditions. For all the treatments, light, temperature and nutrition providing conditions remained same. The seedlings were grown in hydroponic containers. For the cold treatment, plants were incubated at 4°C under continuous cool white fluorescent illumination at approximately 100 μmol m⁻² s⁻¹ light intensity. For salt, ABA and Osmotic treatment, the plants were watered with 100 mM NaCl, 100 μM ABA and 20% (w/v) PEG 600, respectively (Tang et al. 2013). Samples were harvested at 0, 12, 24 and 48 h, then frozen in liquid nitrogen and stored at −80°C for total RNA extraction.

### Molecular cloning and identification of CSP gene family members in Pak-choi and Chinese cabbage

AtCSP1, AtCSP2, AtCSP3, and AtCSP4 were requested from TAIR from a scan of full-length sequences (BRAD). BrCSPs were surveyed on the whole genome of the Chinese cabbage, were identified and named BrSCP1-5. BrCSP1-6 were cloned using full-length primers and cDNA from root and shoot tissue from stress-treated Pak-choi (Supplementary Table S1).

### Genome-wide characterization and annotation of CSP genes in Pak-choi and Chinese cabbage

To identify the CSPs in Chinese cabbage genome, we collected the available sequence data from Chinese cabbage chiflu genome website (BRAD, http://brassicadb.org/brad/) including 41020 protein coding genes and 283.8 Mb of the genome assembly (version 1.5). An hmmsearch using PFAM domains PF00313 was conducted with default parameters on Chinese cabbage proteomes. All putative CSP homologues were detected in the Pfam and SMART database to confirm whether each of them presents the cold shock domain and the redundant sequence were removed. The corresponding sequence concerning full-length cDNA, DNA, promoter and protein of these BrCSP genes were extracted using Perl script. The information of all the analyzed putative BrCSP homologues is listed in Table 1.

| Name       | GenBank accession /Gene locus | CDS (bp) | Exon | Size (aa) | MW (Kda) | pI | CSD position | Subcellular localization | AT ortholog | E-value |
|------------|-------------------------------|---------|------|----------|----------|----|--------------|--------------------------|-------------|---------|
| BcCSP1     | CabbageG_a_f_g003317          | 822     | 2    | 273      | 27.47    | 7.46| 11–75        | nuclear                  | AtCSP3      | 7.00E-72 |
| BcCSP2     | CabbageG_a_f_g018200          | 772     | 1    | 256      | 25.60    | 7.02| 11–77        | nuclear                  | AtCSP4      | 6.00E-60 |
| BcCSP3     | CabbageG_a_f_g020050          | 619     | 1    | 205      | 19.39    | 6.25| 13–79        | nuclear                  | AtCSP1      | 7.00E-101|
| BcCSP4     | CabbageG_a_f_g0747530         | 504     | 1    | 200      | 18.80    | 5.71| 13–79        | nuclear                  | AtCSP4      | 3.00E-61 |
| BcCSP5     | CabbageG_a_f_g033354          | 550     | 1    | 182      | 17.76    | 6.75| 12–79        | nuclear                  | AtCSP2      | 6.00E-45 |
| BcCSP6     | CabbageG_a_f_g039616          | 550     | 1    | 182      | 17.73    | 6.25| 12–78        | nuclear                  | AtCSP4      | 8.00E-58 |

Notes: bp, base pair; aa, amino acids; CDS, coding sequence; Size, number of amino acids; MW, molecular weight; pI, the isoelectric point; CSD, cold shock-domain; AT, Arabidopsis thaliana. The subcellular localization analysis was performed by WoLFPSORT (Horton et al. 2007) and the k-value used for kNN is 14.

### Protein properties and sequences analyses

In this study, we used DNAMAN software (Lynnon Biosoft) to analyze the deduced amino acid sequence of CSP. To perform multiple sequence alignment of CSP, we performed MEGA 5.0 software by using the default parameters and manual correction (Tamura et al. 2014). MEGA 5.0 software was used to conduct sequence analysis and phylogenetic tree drawings. Bootstrap values were estimated with 1000 replicates. The conserved motifs were analyzed using MEME Suite (http://meme.nbcr.net/meme/) with the default settings except the maximum width was set to 200, and the minimum and maximum numbers of motifs were defined as 2 and 10, respectively (Table 2).

### Sequence alignment and phylogenetic analysis

Phylogenetic analysis and amino acid alignment were performed with MUSCLE in MEGA 5.0 and visualized using Jalview. Multiple sequence alignment of CSP proteins was done using MUSCLE, and the phylogenetic tree was constructed using MEGA 5.0 by the Maximum-likelihood method with 1000 bootstrap replicates. The schematic representation of the conserved motifs in the CSP proteins from Pak-choi, Chinese cabbage and Arabidopsis as detected by MEME analysis and each motif is represented by a color box numbered at the bottom. Supplementary Table S3, listing of stress cis-elements on the promoter region of CSP genes from Pak-choi and Chinese cabbage.

### Relative expression of BcCSP and BrCSP genes family under multiple abiotic stresses

Total RNA from Pak-choi and Chinese cabbage leaves were purified, respectively and reverse transcription and quantitative PCR, were performed as described previously with gene-specific primers in Supplementary Table S2, and GAPDH as the endogenous control. To investigate the changes of BcCSP and BrCSP family genes under multiple abiotic stresses treatments, we performed qRT-PCR. At the five-leaf stage seedlings of Pak-choi (B. rapa ssp. chinensis cv. Suzhouqing) and Chinese cabbage (B. rapa ssp. pekinensis), which were

| BrCSP1     | Bra002099: A07 | 2685 | 13 | 894 | 97.27 | 8.60 | 619–685 | mitochondrial | 10.0 | AtCSP3 | 2.00E-95 |
| BrCSP2     | Bra011655: A01 | 855  | 1  | 284 | 28.49 | 7.76 | 11–78   | nuclear      | 14   | AtCSP4 | 5.00E-81 |
| BrCSP3     | Bra017742: A03 | 792  | 1  | 263 | 26.16 | 7.03 | 11–77   | nuclear      | 13   | AtCSP4 | 2.00E-104|
| BrCSP4     | Bra030325: A04 | 618  | 1  | 205 | 19.39 | 6.25 | 13–79   | nuclear      | 12.0 | AtCSP4 | 6.00E-60 |
| BrCSP5     | Bra031159: A09 | 600  | 1  | 199 | 18.79 | 5.70 | 13–79   | nuclear      | 11.0 | AtCSP4 | 1.00E-62 |

Notes: bp, base pair; aa, amino acids; CDS, coding sequence; Size, number of amino acids; MW, molecular weight; pI, the isoelectric point; CSD, cold shock-domain; AT, Arabidopsis thaliana. The subcellular localization analysis was performed by WoLFPSORT (Horton et al. 2007) and the k-value used for kNN is 14.
Table 2. CSP sequence alignment among Pak-choi, Chinese cabbage and Arabidopsis.

| CSP | Alignment of CSP sequence among Pak-choi, Chinese cabbage and Arabidopsis. | Alignment of CSP sequence among Pak-choi, Chinese cabbage and Arabidopsis. | Alignment of CSP sequence among Pak-choi, Chinese cabbage and Arabidopsis. | Alignment of CSP sequence among Pak-choi, Chinese cabbage and Arabidopsis. |
|-----|----------------------------------------------------------------------------|----------------------------------------------------------------------------|----------------------------------------------------------------------------|----------------------------------------------------------------------------|
| AtCSP3 | MAMEMDAASRGGSGVSGDSDKGCGDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MAMEMDAASRGGSGVSGDSDKGCGDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MAMEMDAASRGGSGVSGDSDKGCGDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MAMEMDAASRGGSGVSGDSDKGCGDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |
| BcSCD6 | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |
| BcSCD5 | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |
| BcSCD3 | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |
| BcSCD2 | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |
| BcSCD1 | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |
| BrCSP4 | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |
| BrCSP3 | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |
| BrCSP2 | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |
| BrCSP1 | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG | MSGGGDVNXMSGGDNRKTVKDFIDPDGGIDLDHVSQGIGSSTVGDSRLTGSELSEYSGASDTSKAIETPAGGSGNILKSESNSSGSGGNCFCEGVNHMADKCDGGSGG |

Results and discussion

Identification of CSP gene family in Pak-choi, Chinese cabbage and Arabidopsis

To identify all possible members of the CSP family in Pak-choi, Chinese cabbage and Arabidopsis, we first used Arabidopsis CSP proteins to search against the Chinese cabbage genome and then built a Brassica-specific HMM-profile on the basis of the results of the searches to identify all the members of the CSP family in Chinese cabbage. We found domain-containing genes in the Chinese cabbage genome, named BrCSP1 to BrCSP5 based on their sequence length (Table 1). Chinese cabbage has more CSP domain genes utilized for qRT-PCR analysis. Plants in different treatment groups were prepared separately. For cold stress treatment, seedlings were transferred to 4°C for varying lengths of time (0, 12, 24 and 48 h). For ABA, NaCl and osmotic stresses treatments, seedlings were raised through 2012). Calculating the standard error of the average of three replicates.
than Arabidopsis (4 AtCSP genes), possibly due to its larger genome and genome triplication event. CSPs in Chinese cabbage were easily confirmed with the conserved CSP domain analysis, with the exception of two CSP proteins, BrCSP37 and BrCSP39, which have the VH domain. Besides, we found 6 CSP domain-containing genes in the Pak-choi genome, named BcCSP1 to BcCSP6 based on their sequence length (Table 1).

**Sequence alignment and phylogenetic analysis of CSP proteins**

The deduced amino acid sequence of CSPs was aligned with CSP protein sequences from several representative monocots and dicots by using DNAMAN (Lynnon Biosoft). As is shown in Figure 1, these CSPs were similar to other plant CSPs in sequence, higher sequence similarities were detected in the N-terminus. The sequence alignment showed that the CSPs proteins all have CCHC zine finger, Glycine-rich region and cold shock domain (Figure 1). The results revealed that the BcCSPs protein sequence was more homologous to Chinese cabbage protein sequences than AtCSP protein sequences (Figure 1), suggesting that BcCSP proteins in dicot plants are conserved. More importantly, all CSPs all have same motif 1 to motif 6 (Figure 2(b)), all have same gene module (Supplementary Table S2). As is shown in Figure 2(d), we found that structure of BcCSP1 protein and AtCSP3 protein has a 53.92% similarity, BcCSP2 and AtCSP2 (59.27%), BcCSP3 and AtCSP4 (73.08%), BcCSP4 and AtCSP4 (75.00%), BrCSP3 and AtCSP1 (62.05%), BrCSP5 and AtCSP4 (73.03%), BrCSP6 and AtCSP4 (75.00%), which revealed that BcCSP5 protein and AtCSP4 protein have more high similarity. Furthermore, those results revealed the BcCSPs protein might be involved in low temperature stress.

In order to show the evolution of CSPs protein vividly, we collected the CSP proteins coding sequence of Pak-choi and other species. Then phylogenetic tree was constructed by
Figure 2. Phylogenetic analyses of CSP proteins from ten plant species. Multiple sequence alignment of CSP domains was done using MUSCLE, and the phylogenetic tree was constructed using MEGA 5.0 by the Maximum-likelihood method with 1000 bootstrap replicates. The tree was divided into 2 phylogenetic clades, designated as I and II. Members of Pak-choi, Chinese cabbage and *Arabidopsis*, were clustered into the two terminals of phylogenetic tree, respectively.

Figure 3. Expansion and evolution of Plant CSP proteins. Phylogenetic tree and major domain of CSP proteins. (a) Phylogenetic analysis of 212 CSP proteins from Chinese cabbage (127), *Arabidopsis* (85) showing similar groups in all of the plant species. In total, 10 clades with different colors that were formed by CSP proteins are also marked. (b) The kinases are drawn with the N-terminus to the left and C-terminus to the right. Catalytic domains have been aligned to facilitate comparison. Auto-hibitory domains are shown. EF-hands are shown as green boxes within the regulatory domains. The CSP C-terminus domains are shown. Acidic patch domains are shown. The CSP NAF/FISL domains are shown as dark light.
MEGA 5.0 (Figure 2a). BrCSP2 and BcCSP1, their evolutionary relationship is the closest, with 100%. BrCSP6 and AtCSP2, their evolutionary relationship is the farthest away, with 74%. As is shown in Figure 3, the phylogenetic tree from 12 species was divided into two Groups, named Group I and Group II, respectively. BcCSP1, BcCSP2, BrCSP2 and BrCSP3 in Group I, some CSP homologous protein from Arabidopsis, Brassica oleracea and B. rapa were also divided into this branch, it could be seen that BrCSP2 protein was highly homologous to OsCSP1 from pak-choi. BcCSP4 in Group II, it could be seen that its homology with CSP4 proteins from Pak-choi were higher than that from Oryza sativa and Zea mays. Group II contains BcCSP2, BcCSP3 and BcCSP5 homologous proteins, BcCSP2 and BcCSP5 on the same branch. In general, BcCSP proteins from Pak-choi were more homologous to BrCSP proteins and were highly conservative, then were homologous to CSP members from Arabidopsis and other dicotyledonous plants, and BcCSP evolution is most distant from gramineae homologous proteins. CSP proteins from O. sativa were homologous to CSP proteins from Z. mays, compared with those from other crops. The homology of CSP proteins from Daucus carota and CSP proteins from Solanum lycopersicum were higher than CSP proteins from Glycine max.

In Figure 2, the location of some different CSP proteins motif sites is shown. Each block shows the position and strength of a motif site. As was shown in Figure 2(b), there were at least six motif sites for each CSP protein, motif location of AtCSP2 protein was very similar to motif location of AtCSP3 protein, motif location of AtCSP4 protein was similar to motif location of AtCSP5 protein. Interestingly, motif location of OsCSP1 protein was similar to motif location of ZmCSP1 protein. More importantly, motif location of BcCSP1 protein was very similar to motif location of AtCSP1 protein. Therefore, we hypothesized that BcCSP1 and AtCSP1 had similar functional responses to abiotic stress (Supplementary Figure S2).

CSP orthologs and paralogs in Chinese cabbage and Arabidopsis

In Chinese cabbage and Arabidopsis, OrthoMCL software identified homologous gene pairs. In addition, there are orthologous CSP gene pairs in Chinese cabbage and Arabidopsis, respectively (Figure 4). In homologue pairs in Chinese cabbage and Arabidopsis, we found that AtCSP genes At02 and At04 all have five Chinese cabbage homologues, demonstrating that CSP is replicated with triparylation in Chinese cabbage genome. However, the number of genes in Chinese cabbage genome was significantly less than three times that of Arabidopsis, indicating that the gene was lost in the formation of polyploid species. For example, three AtCSPs do

![Figure 4](image-url) Paralogs Ortholog groups of CSP in Chinese cabbage and Arabidopsis Genome. The lines regarding orthologous gene pairs; Chinese cabbage paralogous gene pairs and Arabidopsis paralogous gene pairs are colored red.
not have homologous genes in Chinese cabbage. However, AtCSP gene At01, At03 and At05 do not have Chinese cabbage homologues. Based on the results of comparative genomic analysis, we can infer that some BrCSP genes function according to their Arabidopsis homologues, which greatly facilitate our research on the role of BcCSP, BrCSP genes in Pak-choi and Chinese cabbage.

**Phylogenetic and taxonomic analysis of CSP gene**

CSP transcription factor both Pak-choi, Chinese cabbage and Arabidopsis is studied by establishing a phylogenetic tree. Information on plant CSP protein sequences is listed in Supplementary Table S4. Six BcCSP genes, five BrCSP genes and four AtCSP genes were used for comparative analysis, and then a rootless phylogenetic tree was constructed. The evolutionary tree is divided into two distinct subfamilies, group I and group II (Figure 3).

To further understand the diverse protein structure, the motif and gene structure of Pak-choi, Chinese cabbage and Arabidopsis CSP were predicted (Figure 5). Genes from the same subfamily share similar motifs and may have similar functions. By comparing Pak-choi, Chinese cabbage and Arabidopsis CSP proteins, we found that the genetic structure of each subfamily is very similar. All CSPs contain MYB-box, which indicates that all CSPs contain highly conserved domains. In addition, AtCSP1, AtCSP2, BcCSP4, and BcCSP5 all have HSE, so we suspect that they may respond to some abiotic stress. The gene structure shown is essentially a mock-compliant analysis, and members of the same class usually have the same gene intron and exon patterns (Figure 5).

**Relative expression of CSP gene family from Pak-choi and Chinese cabbage under abiotic stresses**

The relative transcript levels of BcCSPs and BrCSPs were quantified in seedlings plants grown under multiple stress condition. To investigate the expression patterns of BcCSPs and BrCSPs under different abiotic stresses, we performed qRT-PCR analysis for CSPs gene. Research indicated that CSPs showed various expression patterns under the tested abiotic stresses. Under cold treatment, CSPs expression was marginally down-regulated after 0.5 h and then markedly up-regulated at 2 h after the initiation of exogenous cold treatment, reaching a maximum level of 1.6-fold compared with the control; however, in the salt treatment, the induction dynamics of CSPs at transcript level was rapidly increased and reached a maximum rate at 8 h after treatment, then followed by a slow descent; the CSPs transcript level had a rapid uptake process at 1 h followed by a slow descent under ABA treatment; in cold treatment, the induction dynamics of CSPs was similar to those under ABA treatment (Figure 6).

After the treatment with ABA or cold, CSPs mRNA expression in leaves increased significantly at 1 h, indicating that CSPs was involved in co-responding to ABA and cold treatments. It is unlikely that CSPs is related to recovery from translational arrest under low temperature, because the expression of CSPs was not increased promptly after the drop in temperature but was induced after 1 h of cold stress treatment. As shown in Supplementary Figure S1, all BrCSPs are expressed in roots, stems, and leaves. Simultaneously, the expression level of BrCSP1 in stem is higher than that in root or leaf, the expression level of BrCSP4 in leaf is the lowest compared to root or leaf. In addition, we also tested the expression levels of AtCSP1 to AtCSP4 under ABA, Cold, Salt, and Osmotic treatment conditions, we found that almost all genes respond to these treatments, among them, AtCSP4 responds to all treatments with the highest level of expression (Supplementary Figure S3). Thus, these data suggested that CSPs expression was influenced by ABA, cold, salt and Osmotic treatments and that CSPs was coordinately expressed in response to ABA and cold stimulation in Pak-choi leaves.

**Conclusions**

The cold shock domain (CSD) proteins have been initially identified in Escherichia coli, due to massive induction in...
response to a sudden drop in temperature and thus been named so (Jones et al. 1987), the cold shock domain (CSD) is a nucleic acid binding domain that is widely conserved from bacteria to higher plants and animals (Karlson and Imai 2003). In plants, low-temperature tolerance is induced upon exposure to low and non-freezing temperatures. This phenomenon, known as cold acclimation, is a cumulative process that is activated once temperatures fall below a critical threshold temperature (Fowler 2008). As one of the most popular fresh vegetables, Pak-choi is often subjected to a variety of biological stresses and abiotic stresses throughout all the growth and development stages (Wei et al. 2016). Low temperature limits the normal growth development of plants directly through the inhibition of metabolic reactions, limiting the normal growth development of plants indirectly through cold-induced osmotic and oxidative stresses (Chinnusamy et al. 2007). In this study, CSPs gene was isolated from Pak-choi based on amino acid sequence alignment.

Figure 6. qRT-PCR of CSP genes from leaves of Pak-choi and Chinese cabbage in response to ABA, Cold, Salt and Osmotic treatments. x-axis, different CSP genes; y-axis, gene relative qRT-PCR expression values, error bars SD.
The results revealed that the BcCSP1 protein sequence was highly similar to the CSP protein sequences of Arabidopsis and was homologous to the CSP protein sequences of O. sativa, Z. mays and G. max, indicating that the CSP proteins in the dicotyledonous plants was conserved (Figure 1). Phylogenetic and structural analysis of the putative BcCSP1 protein and other CSP proteins available from GenBank and other databases was carried out. The results suggested that these orthologous genes might be derived from a common ancestor and tend to be conserved during evolution.

CSPs expressions are conventionally associated with cold shock acclimation (Herman 1993; Phadtare and Inouye 1999; Gualerzi et al. 2003). qRT-PCR analysis revealed that BcCSPs and BrCSPs played an important regulatory role in response to ABA, Cold, Salt and Osmotic treatments in Pak-choi and Chinese cabbage. However, the expression patterns of BcCSPs and BrCSPs differ from the expression patterns of AtCSPs in some aspects. Under ABA or Osmotic treatment, BcCSP1, BcCSP2, BcCSP3, BcCSP4, BcCSP5 and BcCSP6 have relatively obvious responses, and more importantly, their expression levels show a trend of rising first and then decreasing. Under Cold or Salt treatment, BcCSP1, BcCSP2, BcCSP3, BcCSP4 and BcCSP5 have relatively obvious responses, their expression levels show a trend of rising first and then decreasing, however, there was no significant change in the expression level of BcCSP6 (Figure 6). Under ABA treatment, BrCSP1, BrCSP2, BrCSP3, BrCSP4 and BrCSP5 have relatively obvious responses, their expression levels show a trend of rising first and then decreasing, last rising. Under Cold treatment, BrCSP1 and BrCSP2 have relatively obvious responses, their expression levels show a trend of rising first and then decreasing, last rising, however, there was no significant change in the expression level of BrCSP3, BrCSP4 and BrCSP5. Under Salt treatment, BrCSP1 and BrCSP5 have relatively obvious responses, their expression levels show a trend of rising first and then decreasing, however, there was no significant change in the expression level of BrCSP2, BrCSP3 and BrCSP4. Under Osmotic treatment, BrCSP1, BrCSP2 and BrCSP3 have relatively obvious responses, their expression levels show a trend of rising first and then decreasing, however, there was no significant change in the expression level of BrCSP4 and BrCSP5.

The expression and regulation of AtCSP gene family has been studied, the expression of CSP gene family is positively regulated by the presence of sucrose in the medium. These results suggested that the physiological functions of BcCSPs might differ from those of AtCSPs in response to multiple stresses. BcCSPs might cross-respond to multiple stresses and produce an overlapping effect or cross-protection against environmental stress factors.

A total of 6 BcCSP genes and 5 BrCSP genes were identified from two Brassica crops, and all of these genes were further system investigated. All CSP orthologs were analyzed by phylogenetic analysis, syntenic analysis, genomic duplication and evolutionary divergence. Phylogenetic analysis of the CSP family among Pak-choi, Chinese cabbage and Arabidopsis showed that the conserved organization of the CSP family implied that these genes underwent ancient and recent gene duplication events from a common origin and were retained over a long period of domestication for each genome. In addition, the increased number of CSP genes correlates closely with the frequency of duplication events, and these duplications might contribute to the survival of plants in adverse conditions and the functional diversity of the CSP genes. Further, tissue-expression profiles of the CSP genes showed several BrCSP genes that were highly expressed in different tissues and several BcCSP genes that were co-expressed under multiple stresses patterns, suggesting that they might play more important roles in developmental and metabolic processes. This study shows the CSP proteins conserved evolution and functional diversity in two cultivated Cabbage (Pak-choi and Chinese cabbage) enabled us to further identify stress-responsive CSP genes and species-specific genes in response to multiple stresses conditions.

Acknowledgments

Conception and design of the study: Feiyi Huang and Jun Tang. Performed the experiments: Jun Tang, Feiyi Huang, Jin Wang and Xilin Hou. Wrote the paper: Feiyi Huang, Jin Wang, and Xilin Hou. Manuscript revision and approval: Feiyi Huang, Jin Wang and Xilin Hou. All authors read and approved the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the State Key Program of National Science Foundation of China [grant number 31330067], China Agriculture Research System [CARS-23-A-06], Key Projects of National Key Research and Development plan [2017YFD0101803], National Key R&D Program of China [2016YFD0101701], China Postdoctoral Science Foundation [2018M640493].

Notes on contributors

Feiyi Huang, is a Teacher in Nanjing Agricultural University, Vegetable Science.

Jin Wang is a PhD student in Nanjing Agricultural University, Vegetable Science.

Jun Tang is a Teacher in Nanjing Agricultural University, Vegetable Science.

Xilin Hou is a Professor in Nanjing Agricultural University, Vegetable Science. He is mainly engaged in biomolecular genetic breeding and bioinformatics of non-heading Chinese cabbage, and he has also cultivated some new varieties.

ORCID

Xilin Hou  http://orcid.org/0000-0003-0776-2959

References

Barnes AC, Benning C, Roston RL. 2016. Chloroplast membrane remodeling during freezing stress is accompanied by cytoplasmic acidification activating sensitive to freezing. Plant Physiol. 171:2140–2149.

Chaikam V, Karlson D. 2010. Functional characterization of two cold shock domain proteins from Oryza sativa. Plant Cell Environ. 31:995–1006.

Chinnusamy V, Zhu J, Zhu JK. 2007. Cold stress regulation of gene expression in plants. Trends Plant Sci. 12:444.

Fowler DB. 2008. Cold acclimation threshold induction temperatures in cereals. Crop Sci. 48:1147–1154.

Fusaro AF, Bocca SN, Ramos RLB, Barrôco RM, Magioli C, Jorge VC, Coutinho TC, Rangel-Lima CM, Rycke RD, Inzé D. 2007. AtGRP2, a cold-induced nucleo-cytoplasmic RNA-binding protein, has a role in flower and seed development. Planta. 225:1339–1351.
Gilmour SJ, Zarka DG, Stockinger EJ, Salazar MP, Houghton JM, Thomashow MF. 1998. Low temperature regulation of the Arabidopsis CBF family of AP2 transcriptional activators as an early step in cold-induced COR gene expression – Gilmour – 2001 – The plant Journal – Wiley Online Library. Plant J Cell Molec Biol. 16:433.

Gualerzi CO, Giuliodori AM, Pon CL. 2003. Transcriptional and post-transcriptional control of cold-shock genes. J Mol Biol. 331:527–539.

Herman IM. 1993. Actin isoforms. Curr Opin Cell Biol. 5:48.

Horton P, Park KJ, Obayashi T, Fujita N, Harada H, Adams-Collier CJ, Nakai K. 2007. PSORT: protein localization predictor. Nucleic Acids Research. 35(Web Server):W585–W587.

Jones PG, Vanbogelen RA, Neidhardt FC. 1987. Induction of proteins in response to low temperature in Escherichia coli. J Bacteriol. 169:2092–2095.

Karlson D, Imai R. 2003. Conservation of the cold shock domain protein family in plants. Plant Physiol. 131:12–15.

Karlson D, Nakaminami K, Thompson K, Yang Y, Chaikam V, Mulinti P, Gusta LV, Wisniewski ME, Tanino KK. 2009. Plant cold-shock domain proteins: on the tip of an iceberg [A]. In Gusta LV, Wisniewski ME, Tanino KK, editors. Plant cold hardness from the laboratory to the field [M]. Cambridge: Cabi; p. 43–54.

Li C. 2014. Cold shock domain proteins in normal and leukemic hematopoiesis [dissertations & theses – gradworks].