Mol Genet Genomics (2015) 290:239–255
DOI 10.1007/s00438-014-0912-7

ORIGINAL PAPER

Genome-wide analysis of the MADS-box gene family in *Brassica rapa* (Chinese cabbage)

Weike Duan · Xiaoming Song · Tongkun Liu · Zhinan Huang · Jun Ren · Xilin Hou · Ying Li

Received: 29 June 2014 / Accepted: 28 August 2014 / Published online: 13 September 2014

Abstract The MADS-box gene family is an ancient and well-studied transcription factor family that functions in almost every developmental process in plants. There are a number of reports about the MADS-box family in different plant species, but systematic analysis of the MADS-box transcription factor family in *Brassica rapa* (Chinese cabbage) is still lacking. In this study, 160 MADS-box transcription factors were identified from the entire Chinese cabbage genome and compared with the MADS-box factors from 21 other representative plant species. A detailed list of MADS proteins from these 22 species was sorted. Phylogenetic analysis of the *BrMADS* genes, together with their *Arabidopsis* and rice counterparts, showed that the *BrMADS* genes were categorised into type I (Mα, Mβ, Mγ) and type II (MIKC^C, MIKC*) groups, and the MIKC^C proteins were further divided into 13 subfamilies. The Chinese cabbage type II group has 95 members, which is twice as much as the *Arabidopsis* type II group, indicating that the Chinese cabbage type II genes have been retained more frequently than the type I genes. Finally, RNA-seq transcriptome data and quantitative real-time PCR analysis revealed that *BrMADS* genes are expressed in a tissue-specific manner similar to *Arabidopsis*. Interestingly, a number of *BrMIKC* genes showed responses to different abiotic stress treatments, suggesting a function for some of the genes in these processes as well. Taken together, the characterization of the *B. rapa* MADS-box family presented here, will certainly help in the selection of appropriate candidate genes and further facilitate functional studies in Chinese cabbage.

Keywords Abiotic stress · Chinese cabbage · Genome-wide analysis · MADS-box transcription factor · qRT-PCR

Communicated by S. Hohmann.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-014-0912-7) contains supplementary material, which is available to authorized users.

W. Duan · X. Song · T. Liu · Z. Huang · J. Ren · X. Hou · Y. Li (✉)
State Key Laboratory of Crop Genetics and Germplasm Enhancement, Ministry of Agriculture, Nanjing Agricultural University, Nanjing 210095, People's Republic of China
e-mail: yingli@njau.edu.cn

W. Duan · X. Song · T. Liu · Z. Huang · J. Ren · X. Hou · Y. Li
Key laboratory of Biology and Germplasm Enhancement of Crops in East China, Ministry of Agriculture, Nanjing Agricultural University, Nanjing 210095, People's Republic of China

Introduction

MADS-box genes encode transcription factors that are involved in developmental control and signal transduction in eukaryotes (Riechmann and Meyerowitz 1997). These genes are found in fungi (Passmore et al. 1988), animals (Norman et al. 1988) and plants (Sommer et al. 1990; Yanofsky et al. 1990). They constitute a large gene family, which is named after a few of its earliest members, *MCM1* (from yeast) (Passmore et al. 1988), *AGAMOUS* (from *A. thaliana*) (Yanofsky et al. 1990), *DEFICIENS* (from *Antirrhinum majus*) (Sommer et al. 1990) and *SRF* (from *Homo sapiens*) (Norman et al. 1988). Previous studies of the MADS-box genes have included a thorough comparison analysis of their roles in plant growth and development. However, there are relatively few analyses of the response of these genes to stress conditions. *Brassica rapa* ssp. *pekinensis* (Chinese cabbage) is one of the subspecies of *Brassica rapa*. This subspecies, which originated in China, is one of the most economically significant vegetable crops

in Asia. Moreover, Chinese cabbage has become a vegetable that is grown worldwide due to its high yield and good quality. Thus, the growth, development and flowering time of this plant are significant for its yield. Recently, the Chinese cabbage (Chiifu-401-42) genome has been sequenced, and this sequence can help us with the analysis of MADS-box genes from the entire genome (Wang et al. 2011). This genome has undergone triplication events since its divergence from *Arabidopsis* (13–17 mya) (Wang et al. 2011); however, a high degree of sequence similarity and conserved genome structure remain between these two species, these traits make *B. rapa* a good species to use to study the retention and ortholog groups of MADS-box genes during genome duplication events. Furthermore, plant growth and development are influenced greatly by numerous plant growth regulators and environmental factors.

MADS proteins are characterised by the presence of a conserved 58–60 amino acids long DNA-binding domain in the N-terminal region, which is known as the MADS domain, and which binds to CArG boxes (Yanofsky et al. 1990). Based on the phylogenetic analysis, the plant MADS gene family is divided into two large lineages, type I and type II, which were generated by an ancestral gene duplication event (Alvarez-Buylla et al. 2000; Becker and Theißen 2003). The type I genes encode SRF-like domain proteins, whereas type II genes encode MEF2-like proteins (De Bodt et al. 2003). The plant type II proteins are named MIKC due to their four domains. In addition to the MADS (M) domain, MIKC type proteins contain the I (intervening), K (keratin-like) and C (C-terminal) domains (Cho et al. 1999). The I domain contributes to dimer formation (Henschel et al. 2002). The K domain is characterised by a coiled-coil structure, which primarily regulates to the dimerisation of MADS proteins (Díaz-Riquelme et al. 2009). The C domain functions in transcriptional activation and in the formation of higher order protein complexes (Honma and Goto 2001). MIKC-type genes have been further divided into two subgroups, $MIKC^C$ and MIKC*, based on sequence divergence at the I domain (Henschel et al. 2002). The MIKC* genes encode proteins that tend to have longer I domains and have a duplicated K domain. The type I lineage groups genes with a relatively simple gene structure (only with one or two exons) that lack the K domain and that have common ancestors. The type I genes are subdivided into three groups, Mα, Mβ, Mγ, based on the sequence of the MADS domain and on the presence of additional motifs. The function of the type I genes appears to be restricted to female gametophyte (*AGL80* and *AGL61*) and seed development (*PHE1*, *PHE2*, *AGL23*, *AGL28*, *AGL40*, *AGL62*) (Köhler et al. 2003; Bemer et al. 2010; Colombo et al. 2008; Masiero et al. 2011).

Plant MIKC genes were first identified as floral organ identity genes in *Antirrhinum majus* and in *Arabidopsis* (Sommer et al. 1990; Yanofsky et al. 1990). Biologists have made great progress in elucidating the roles of these genes in plant development. Further genetic and molecular analyses regarding their biological functions have focused on flower organogenesis, which acts as the major component in the well-known ABCDE model: sepals (A + E), petals (A + B + E), stamens (B + C + E), carpels (C + E), and ovules (D + E) (Zahn et al. 2006). Briefly, a previous study of *Arabidopsis* MIKC genes classified these genes into five functional classes as follows: Class A includes *APETALA1* (*AP1*); class B includes *PISTILATA* (*PI*) and *AP3*; class C includes *AGAMOUS (AG)*; class D includes *SEEDSTICK/AGAMOUS-LIKE11* (*STK/AGL11*); and class E includes *SEPALLATA* (*SEP1, SEP2, SEP3, and SEP4*) (Pinyopich et al. 2003). Other MIKC genes were later identified as being involved in different regulatory steps, such as: (1) Determination of flowering time genes, which include *Suppressor of Overexpression Of Constans1* (*SOC1*) (Samach et al. 2000; Moon et al. 2003a, b), *AGAMOUS-LIKE GENE 24* (*AGL24*) (Liu et al. 2008), *Short Vegetative Phase* (*SVP*) (Lee et al. 2007), *MADS Affecting Flowering* (*MAF1/FLM*), *Flowering Locus c*(*FLC*) (Michaels and Amasino 1999; Ratcliffe et al. 2003) and *AGL15*, *AGL18* (Adamczyk et al. 2007); (2) Fruit ripening genes, which include *SHATTERPROOF 1–2* (*SHP1, SHP2*) and *FUL* (Liljegren et al. 2000); (3) Seed pigmentation and embryo development genes, which include *TRANSPARENT TESTA16* (*TT16*) (Nesi et al. 2002). Apart from reproductive development, MIKC genes also function in vegetative development and root development, such as *AGL12* and *AGL17* genes (Tapia-López et al. 2008).

Some $MIKC^C$ genes have already been shown to play key roles to control flowing time in *Brassica*, such as *BrFLC1*, *2*, *3*, *BcFLC*, *BrAGL20* and *BnAP3* (Pylatuik et al. 2003; Hong et al. 2012; Liu et al. 2013). For example, the overexpression of *BrAGL20* can significantly affect the flowering time of *B. napus*, and *BrFLC* genes act similar to *AtFLC*, with lower expression in early-flowering Chinese cabbage (Hong et al. 2012). Furthermore, plant growth and development are influenced greatly by numerous plant growth regulators and environmental factors. Gibberellin (GA) promotes flower formation and flowering time in biennial plants. Its involvement in flower initiation in plants is well-established, and there is growing insight into the mechanisms by which floral induction is achieved (Mutasa-Göttgens and Hedden 2009). Salicylic acid (SA) also regulates flowering time because SA-deficient plants are late flowering (Martínez et al. 2004). Abscisic acid (ABA) regulates many aspects of plant growth and development (Bezerra et al. 2004; Wilmowicz et al. 2008). As important environmental stress factors, cold and heat also regulate plant growth and development. To learn more about the response of *B. rapa* MADS-box genes to abiotic stresses, we selected these five treatments to explore in this study.

Flower development is controlled by a complex network of interactions between transcription factors, most of them belonging to the MADS-box family (Airoldi and Davies 2012). To get a better picture about the size and phylogeny of the MADS-box family in plants, we sorted and compared the MADS-box genes from 22 different plant species. To better understand these transcription factors in Chinese cabbage, we determined 160 MADS-box genes and analysed the phylogenetic relationships, conserved motifs, retention and ortholog groups between these Chinese cabbage MADS-box genes and *Arabidopsis* MADS-box genes. We further studied the chromosomal locations, gene duplication and tissue-specific expression of *BrMADS* genes. The expression of all of the *BrMIKC*^{C} genes was also investigated under different treatments, which included GA, SA, ABA, heat and cold.

Materials and methods

Identification of MADS-box gene family in Chinese cabbage

All the files that are related to *Brassica* genome sequence data that were used for the identification and annotation of MADS proteins were downloaded from the *Brassica* database (BRAD; http://brassicadb.org/brad/) (Wang et al. 2011). Proteins with SRF-TF domains (PF00319) were retrieved from the Pfam 27.0 database (http://Pfam.sanger.ac.uk/) (Punta et al. 2012). The hidden Markov model (HMM) was used to identify the putative MADS proteins in Chinese cabbage (Finn et al. 2011). To obtain the proteins, first we used the tool hmmsearch, with an expected value (e-value) cut-off 1.0. Then, we verified these sequences using the tool SMART (http://smart.embl-heidelberg.de/) (Letunic et al. 2012), the Pfam database (http://Pfam.sanger.ac.uk/) and the NCBI database (http://www.ncbi.nlm.nih.gov/).

Sequence retrieval

The *Arabidopsis thaliana* MADS proteins were retrieved from the TAIR database (http://www.arabidopsis.org/) according to a previous report by Parenicova et al. (2003). The dataset of predicted *Oryza sativa* MADS proteins was retrieved from previous analyses by Arora et al. (Arora et al. 2007). A MADS-box domain was not found in *LOC_Os02g01360* (OsMADS60), *LOC_Os12g31010* (OsMADS67), and *LOC_Os08g20460* (OsMADS69). The MADS proteins of *Populus trichocarpa*, *Medicago truncatula*, *Glycine max*, *Cucumls sativus*, *Citrus sinensis*, *Citrus clementine*, *Vitis vinifra*, *Sorghum bicolor*, *Zea mays*, *Selaginella moellendorffi* and *Physcomitrella paters* were retrieved from a previous report.

The Pfam database (http://pfam.sanger.ac.uk/) was used to screen the genome assemblies of *Prunus persica*, *Arabidopsis lyrata*, *Capsella rubella*, *Thellungiella halophila*, *Solanum tuberosum*, *Solanum lycopersicum*, *Aquilegia coerulea* and *Volvox carteri*. The genome data were downloaded from the genome browser phytozome (http://www.phytozome.net/), and the evolutionary relationships of these species were determined using the genome browser phytozome and the public database PGDD (http://chibba.agtec.uga.edu/duplication/) (Lee et al. 2013).

Phylogenetic analysis

In the phylogenetic tree, the *Arabidopsis* MADS proteins were used to classify the Chinese cabbage MADS proteins into different groups. Full-length sequences of MADS proteins of Chinese cabbage and *Arabidopsis* were aligned using the ClustalW2 program with default parameters (Thompson et al. 1997). Then, a phylogenetic tree was then constructed by the neighbour-joining method, and bootstrap values were calculated with 1,000 replications using MEGA5.2 (Tamura et al. 2011). Additionally, an *Arabidopsis* MADS proteins phylogenetic tree was used to detect the reliability of this method, and to test and verify the classification, a phylogenetic tree of Chinese cabbage, *Arabidopsis*, rice and grapevine was built.

To estimate the nucleotide divergence between sequences, all nucleotide sequences of Chinese cabbage MADS-box genes were also analysed by MEGA5.2 using the Jukes-Cantor model. Bootstrap (1,000 replicates) analyses were also performed for this estimation.

Identification of conserved motifs and gene structure

To identify the conserved motifs in full-length Chinese cabbage and *Arabidopsis* MADS proteins, the Multiple Expectation-maximisations for Motif Elicitation (MEME) program version 4.9.0 (Bailey et al. 2009) was used with default parameters, except for the following parameters: (1) optimum motif width was set to $\geq$10 and $\leq$100; and (2) the maximum number of motifs was set to identify 15 motifs. The MEME motifs were annotated using the SMART program (http://smart.embl-heidelberg.de) and the Pfam database.

The coding domain sequences (CDS) and DNA sequences of Chinese cabbage MADS-box genes were used to reveal the gene structure using the tool GSDS (http://gsds.cbi.pku.edu.cn/).

Ortholog groups of MADS-box genes in *Brassica* and *Arabidopsis* genome

The program OrthoMCL (http://www.orthomcl.org/cgi-bin/OrthoMclWeb.cgi) (Li et al. 2003) was used to identify

the homologous genes of MADS-box between Chinese cabbage and *Arabidopsis*. Briefly, the tools BLASTP, with an e-value ≤1e−10, and orthomclPairs were applied to find orthologs, inparalogs and coorthologs in these two species. To link these genes to chromosomes, a tool called Circos (Krzywinski et al. 2009) was used. In addition, the Cytoscape software was applied to build the network of these relationships (Shannon et al. 2003).

Chromosome localisation and gene duplications

To determine the physical locations of MADS-box genes, the starting and ending positions of all MADS-box genes on each chromosome were obtained from the BRAD database. The Perl in-house program was used to draw the location images of the Chinese cabbage MADS-box genes. The positions of each Chinese cabbage MADS-box gene on the blocks were verified by searching for homologous genes between *Arabidopsis* and three *B. rapa* subgenomes, including least fractionated (LF), medium fractionated (MF1) and most fractionated (MF2) genomes (http://brassicadb.org/brad/searchSynteny.php) (Wang et al. 2011; Cheng et al. 2013).

To determine the gene duplications, first, the CDS sequences of Chinese cabbage MADS-box genes were blasted against each other (evalue <1e−10, identity >85 %), and then Ks values were calculated for all pair-wise alignments of these genes, which previously obtained by blast, using the method of Nei and Gojobori as implemented in KaKs_calculator (Zhang et al. 2006). Lastly, based on phylogenies, the nucleotide divergence (Dist <0.1) was used as the final standard (Lynch and Conery 2000). The purple lines were used to link the duplicate genes on different chromosomes.

Chinese cabbage RNA-seq data analysis

For the expression profiling of Chinese MADS-box genes, we utilised the Illumina RNA-seq data that were previously generated and analysed by Tong et al. (2013). Six tissues of *B. rapa* accession Chiifu-401-42, including callus, root, stem, leaf, flower, and silique, were analysed. Two samples of root and leaf tissues were generated from different batches of plants. The transcript abundance is expressed as fragments per kilobase of exon model per million mapped reads (FPKM) values. Heat maps for Chinese cabbage MADS-box genes were generated, which have positive FPKM values in at least one or more of the samples.

Plant material and treatments

The Chinese cabbage cultivar Chiifu-401-42 was used for the experiments. Whole genome sequencing has been completed for this cultivar; thus, this cultivar is a typical cultivar for Chinese cabbage research. Seeds were grown in pots containing a soil: vermiculite mixture (3:1) in the greenhouse of Nanjing Agricultural University, and the controlled-environment growth chamber programmed is light 16 h/25 °C, dark 8 h/20 °C (Song et al. 2013). One month later, seedlings at the five-leaf stage were transferred to growth chambers that were set at 4 or 38 °C under identical light intensity and day length as the cold and heat treatments. Simultaneously, for acclimation, some plants were cultured in 1/2 Hoagland's solution in plastic containers, with the pH at 6.5 (Jensen and Bassham 1966). After 5 days of acclimatisation, plants were cultured in the following four treatments: (1) Control; (2) 100 μM ABA; (3) 100 μM GA; (4) 100 μM SA. At 4 and 12 h after treatment, the young leaf samples were collected, frozen in liquid nitrogen and stored at −70 °C for further analysis.

RNA isolation and quantitative real-time PCR

Total RNA was isolated from 100 mg of frozen tissue using an RNA kit (RNAsimply Total RNA Kit, Tiangen, Beijing, China) according to the manufacturer's instructions. Five micrograms of each sample were reverse transcribed into cDNA using the PrimeScript RT reagent Kit (TaKaRa). The specific primers of Chinese cabbage MADS-box genes and the housekeeping *actin* gene (Bra028615) were designed using the Primer Premier 5.0 software (Supplementary Table 11). To verify the primer specificity, we used the program BLAST against the Chinese cabbage genome. The qRT-PCR assays were performed with three biological and three technical replicates. Each reaction was performed in a 20 μL reaction mixture containing a diluted cDNA sample as the template, 2× Power SYBR Green PCR Master Mix (Applied Biosystems), and 400 nM each of forward and reverse gene-specific primers. The reactions were performed using a MyiQ Single-Color Real-Time PCR Detection System (Bio-Rad, Hercules, CA) with the following cycling profile: 94 °C for 30 s, followed by 40 cycles at 94 °C for 10 s, and 58 °C for 30 s. A melting curve (61 cycles at 65 °C for 10 s) was generated to verify the specificity of the amplification (Song et al. 2013). The relative expression ratio of each gene was calculated using the comparative C_t value method (Heid et al. 1996). The MADS-box gene expression cluster from each stress treatment was analysed using the Cluster program (http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm) (Eisen et al. 1998), and the results were shown using the TreeView software (http://jtreeview.sourceforge.net/).

Results

Identification and classification of MADS-box genes in Chinese cabbage and comparative analyses

To identify the putative MADS proteins in the Chinese cabbage genome, a HMM search resulted in the identification of 164 proteins. Subsequently, all 164 protein sequences were subjected to Pfam and SMART analyses, which resulted in the identification of 162 MADS proteins, called BrMADS001 to BrMADS162 according to the hmmsearch e-value (Supplementary Table 1). Simultaneously, by performing a homology search against *Arabidopsis* and by analyzing the gene structure, two genes were removed. *BrMADS047* and *BrMADS124* contained other functional domains, while their homologs were non-MADS genes (Supplementary Fig. 1).

To pre-classify the Chinese MADS-box genes, a phylogenetic relationship with *Arabidopsis* MADS-box genes was built (Supplementary Fig. 2). In total, 95 genes were determined to be type II MADS-box genes (including MIKC^c and MIKC*), with twofold more members than that in *Arabidopsis.* However, 65 of these genes were confirmed to be type I MADS-box genes (including the Mα, Mβ and Mγ groups), which is comparable to that in *Arabidopsis*.

To perform comparative genomic analyses, we searched for MADS protein-coding sequences in the genomes of 22 other plant species. Some of these genes have been published previously, while others are described in this work for the first time (Supplementary Tables 2 and 3). The evolutionary relationships of the species and the number of MADS-box genes in their genomes are shown in Fig. 1. The data that are coloured green were for the first time analysed in this work. The pre-classified groups of these species were based on their phylogenetic relationships with *Arabidopsis* MADS-box genes. The data show that the number of MADS-box genes in Alga, Bryophyta and Pteridophyta is less than that in Angiospermae. Since several whole genome duplication (WGD) events happened during angiosperm evolution, it is likely that this higher number is caused by an elevated duplication frequency, in combination with an increased retention of MADS-box genes that were subjected to neofunctionalization and gained important functions in angiosperm flower development (Doebley and Lukens 1998; Theissen et al. 2000; Nam et al. 2003).

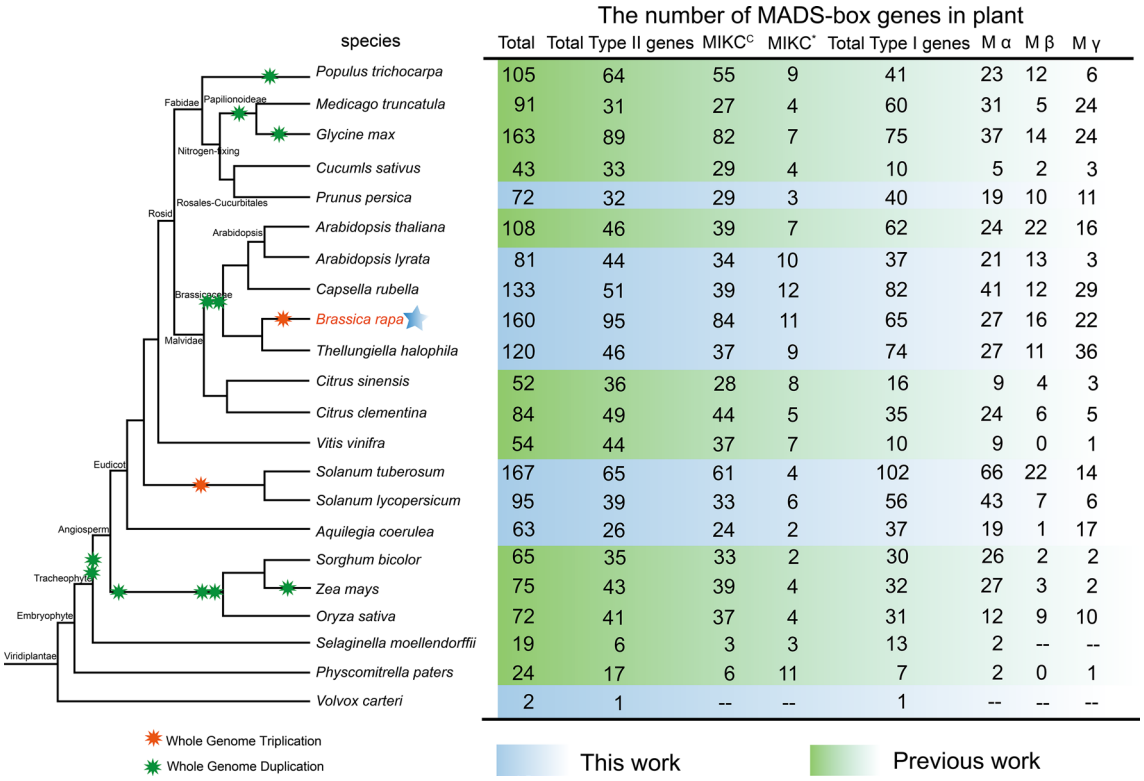

Fig. 1 The evolutionary relationships of the species and the *number* detail of the MADS-box family of each species. The *left* of this figure shows the evolutionary relationships of the species; the *right* of this figure shows the *number* detail of the MADS-box family of each species. The data that are coloured *blue* were described in this work, and the data that are coloured *green* were published in previous works (colour figure online)

Copy number variation and differential retention of MADS-box genes in Chinese cabbage

A comparison of the homologous MADS-box genes in *Arabidopsis* and the three *B. rapa* subgenomes (LF, MF1 and MF2) using the BRAD database revealed that most *BrMADSs* on the conserved collinear blocks have been well-conserved throughout the divergent evolution of *Arabidopsis* and *B. rapa* (Cheng et al. 2012) (Supplementary Table 4). The gene dosage hypothesis predicts that genes whose products are dose-sensitive, interacting either with other proteins or in networks, should be overretained (Thomas et al. 2006; Birchler and Veitia 2007). The type II proteins have been shown to function in large complexes during flower development, while it is still unclear how the type I proteins perform their functions. Interestingly, type II genes have been retained after triplication and fractionation in *B. rapa* at a significantly higher rate than the type I genes (Supplementary Fig. 3a). Most (74 %) type II genes were retained in two or three copies, which is significantly greater than the retention of type I genes (15 %) (Supplementary Fig. 3a), while more (65 %) of the type I genes were completely lost. The proportion of homoeologs retained varied among the three sub-genomes (Supplementary Fig. 3b). In the LF sub-genome, more MADS-box gene homoeologs were retained than other two sub-genomes. The retention of type II genes homoeologs among the sub-genomes was more than that of the type I genes (Supplementary Fig. 3b).

Phylogenetic and classification analysis of *BrMADS* genes

To examine the phylogenetic relationships between *BrMADS* genes in detail, independent phylogenetic trees were constructed with *Arabidopsis* and rice type I and type II proteins (Supplementary Fig. 4 and Fig. 2). The type I proteins were divided into three subfamilies Mα (27), Mβ (16), Mγ (22), whereas the type II proteins were divided into 13 subgroups (Supplementary Table 5). Subgroup TM3-like (*SOC1*) consisted of the highest (16) number of BrMADS type II proteins, whereas subgroup AGL12, AGL6 and Bs (*TT16*) had the lowest members, with only three. Other subgroups contained from four to ten members (Supplementary Fig. 3c). In addition, in the type II group, there are eleven genes that were identified as MIKC*-type.

Finally, we visualized the phylogenetic relationship of the BrMADS proteins with the *Arabidopsis*, rice, soybean and grapevine MADS proteins by building an unrooted tree of the full-length MADS protein sequences. The phylogenetic tree divided these proteins into 5 distinct subfamilies ($MIKC^C$, MIKC*, Mα, Mβ, Mγ) (Supplementary Fig. 2c).

Identification of conserved motifs and gene structure

To compare the differences in the protein structure, MEME was used to identify the conserved motifs among the Chinese cabbage and *Arabidopsis* MADS proteins. The type I and type II MADS proteins of these two species were compared in separate analyses, and for each comparison, fifteen conserved motifs, named motif 1 to motif 15, were identified (Supplementary Fig. 5 and Fig. 3). In general, the MADS proteins were clustered in the same subgroups and shared similar motif composition, which indicates functional similarities among members of the same subgroup (Parenicová et al. 2003). The *Arabidopsis* and Chinese cabbage MADS proteins were found to have similar structure for every subgroup in type II except *BrMADS031, 060, 112* and *103* which with incomplete domains. However, in type I, the protein structure was divergent (Supplementary Fig. 5). This finding indicates that the C-terminal part of the MADS domain in the Mα, Mβ and Mγ groups is more divergent than that in the MIKC group. In type I, except the MADS domain, each of the groups shows a different motif profile, and none of these motifs can be annotated using the tool SMART. The protein motifs shared by the *Arabidopsis* and *Brassica* type I proteins within a clade, show that there is also conservation beyond the MADS domain, although proteins of one clade sometimes show some variation in the motif profile, like for example: *BrMADS118, 128, 144, 136, 157* in Mγ and *BrMADS106, 108, 113* in Mβ (Supplementary Fig. 5).

Simultaneously, the protein structure of BrMADS was analysed using the program MEME. As expected, the commonly shared motifs tend to be in the same group. The motifs were detected by the tool SMART (Supplementary Fig. 6). It will be interesting to characterise the functions of the common motifs within the newly designated groups in relation to the functions of these genes.

In addition to the protein structure, the gene structure was also analysed. We found that all type II *BrMADS* genes have at least three exons, while the number of exons in the type I genes is at maximum two consistent with *AtMADS* genes (Parenicová et al. 2003). Furthermore the first exon (approximately 180 bp) of type II genes conservatively codes the MADS domain. Supplementary Fig. 6 gives an overview of the structures of the Chinese cabbage MADS genes and proteins.

Ortholog groups, chromosomal localization and gene duplication of MADS-box genes

Most angiosperm plant lineages have experienced one or more rounds of ancient polyploidy (Lee et al. 2013). Chinese cabbage has undergone genome triplication since its divergence from *Arabidopsis* (Wang et al. 2011). Generally,

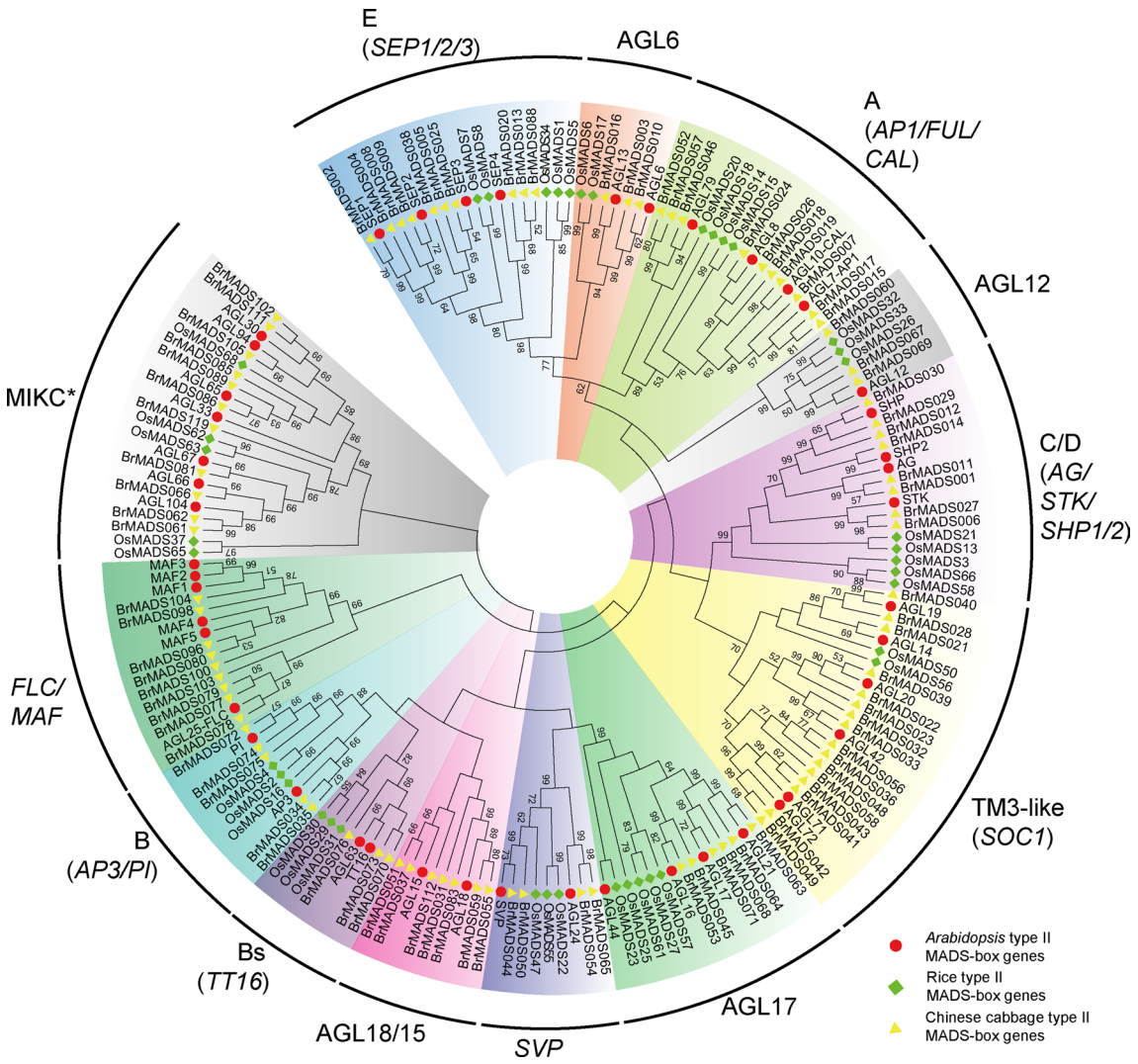

Fig. 2 Phylogenetic tree of Chinese cabbage, *Arabidopsis* and rice type II MADS-box proteins. Phylogenetic analysis of 182 type II MADS proteins from Chinese cabbage (95), Arabidopsis (46) and rice (41) showing similar groups in all of the plant species. In total, 13 clades with *different colours* that were formed by type II MADS proteins are also marked (colour figure online)

the gene number in the Chinese cabbage genome was notably less than three times the *Arabidopsis* gene number because some genes were lost during polyploidy speciation. Additionally, both segmental and tandem gene duplications have significant impacts on the expansion and evolution of gene families in plant genomes. In this study, we analysed the ortholog groups between Chinese cabbage and *Arabidopsis* MADS-box genes using the OrthoMCL program. Then, we identified 67 orthologous gene pairs and 120 co-orthologous gene pairs in the MADS proteins of these two species (Supplementary Table 6). Their visualisation was performed using the Circos software (Fig. 4). Among the orthologous gene pairs of Chinese cabbage and *Arabidopsis*, we found more Chinese cabbage MADS-box homologous genes in *Arabidopsis* chromosome 5 and chromosome 1 than in other chromosomes. To further obtain insight into the correlation of the MADS-box genes in Chinese cabbage and *Arabidopsis,* the networks of MADS-box genes were constructed using these two species orthologous (Supplementary Fig. 7). Among the orthologous gene pairs of Chinese cabbage and *Arabidopsis*, 16 *Arabidopsis* MADS-box genes were found no ortholog with Chinese cabbage MADS-box genes, these genes have been duplicated in *Arabidopsis* after the split. Fifty *Arabidopsis* MADS-box genes have only one ortholog in Chinese cabbage, these genes were present before the split, but two of the three copies have been lost after the *B. rapa* genome triplication (Supplementary Fig. 7a), and 42 *Arabidopsis* genes have co-orthologs in Chinese cabbage, these genes were preferentially retained after the triplication (Supplementary Fig. 7b, c and d). Meanwhile, we found 71 and 60 in paralogous gene pairs in Arabidopsis and Chinese cabbage,

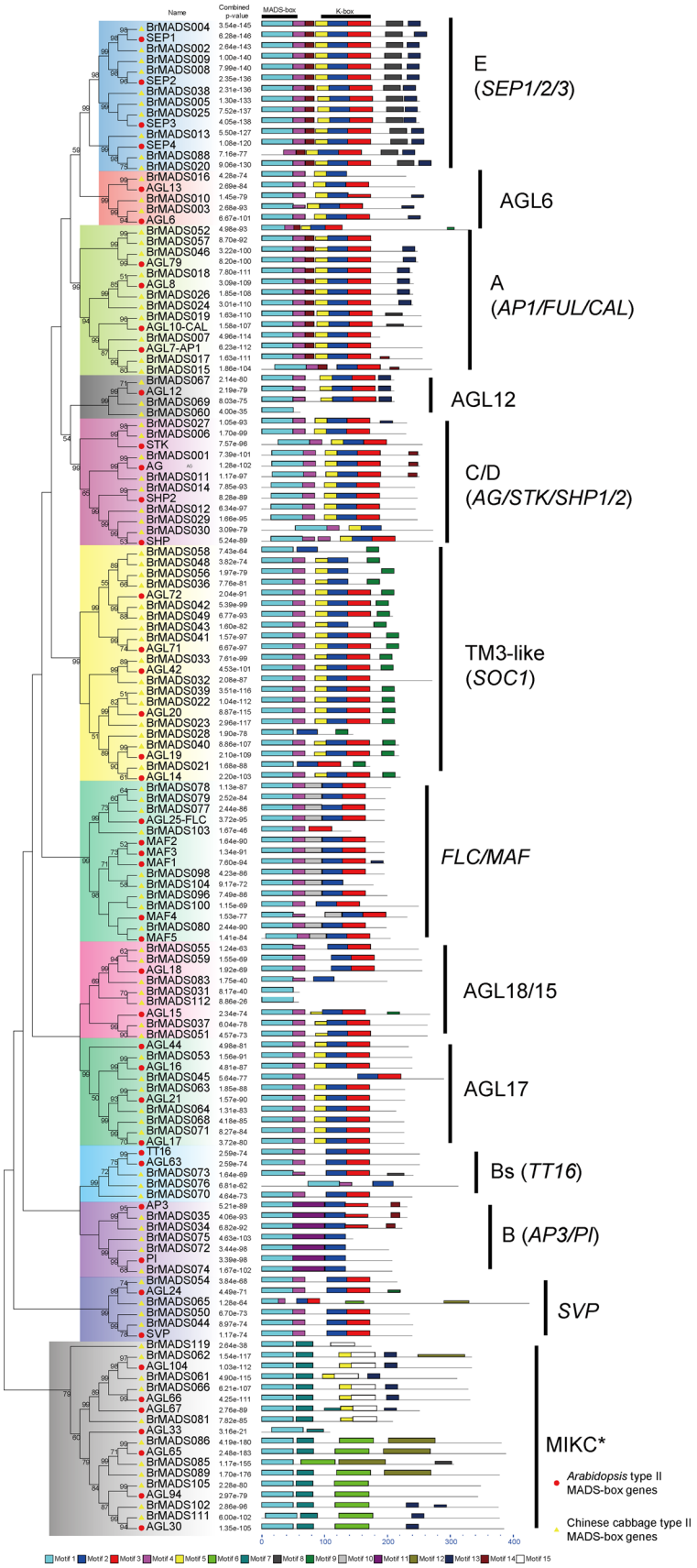

Fig. 3 Phylogenetic relationships and conserved motif compositions of Chinese cabbage and *Arabidopsis* type II MADS proteins. The neighbour-joining tree of Chinese cabbage and *Arabidopsis* type II MADS-box genes and their motif locations

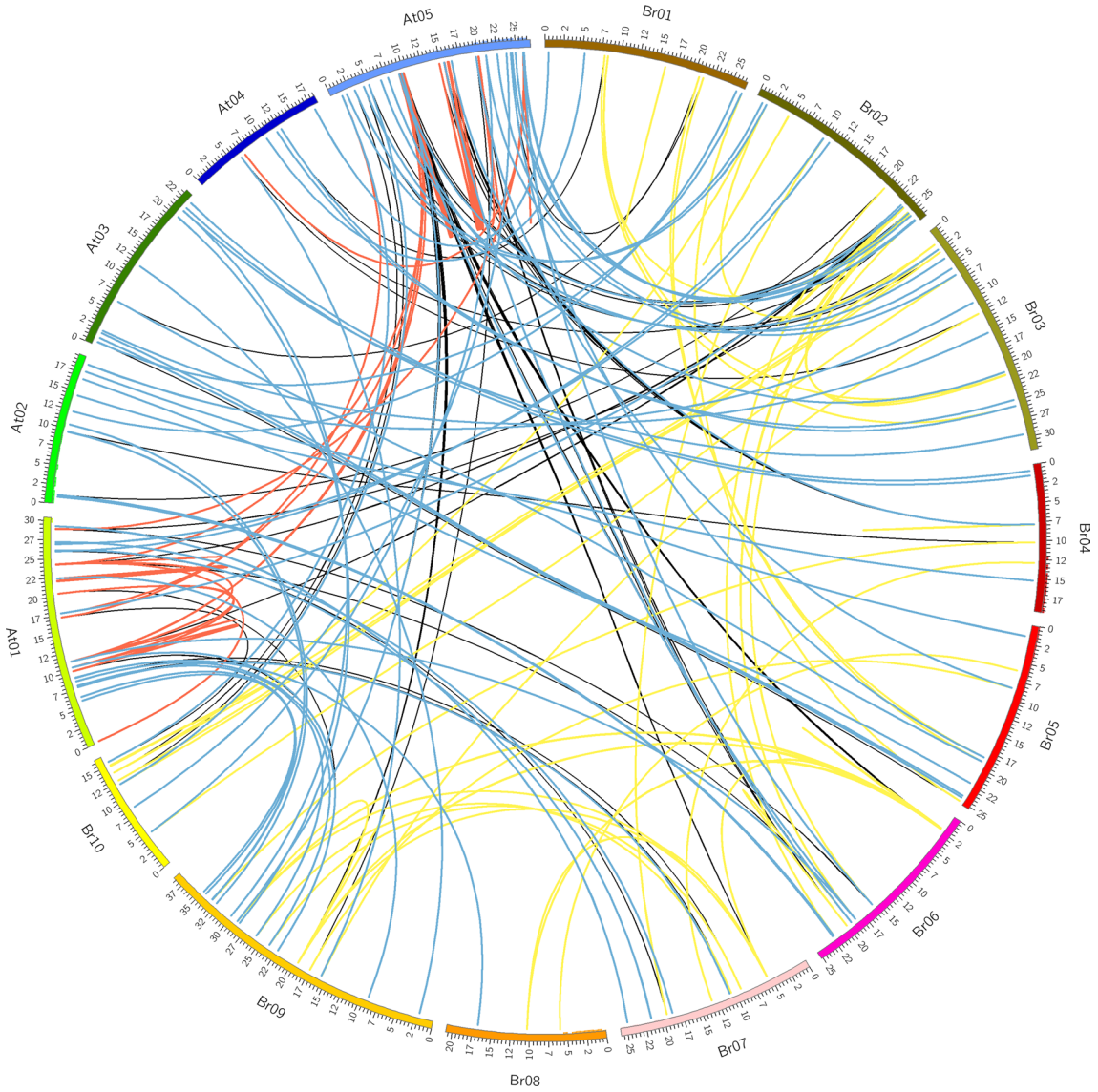

Fig. 4 Ortholog groups of MADS-box genes in *B. rapa* and *Arabidopsis* Genome. Ten Chinese cabbage chromosomes and five *Arabidopsis* chromosomes are coloured different random colours with their names on the periphery. The *lines* in the figure represent four pairs. The *lines* regarding orthologous gene pairs are coloured *blue*; co-orthologous gene pairs are coloured *black*; Chinese cabbage paralogous gene pairs are coloured *yellow* and *Arabidopsis* paralogous gene pairs are coloured *red*. The figure was created using the software Circos (colour figure online)

respectively (Supplementary Table 5 and Fig. 4). From this analysis, we found gene duplication events after the divergence of Chinese cabbage and *Arabidopsis* resulted in a high number of paralogous and co-orthologous genes in both species. While in *Arabidopsis* the type I subfamily has predominantly expanded, it is in Chinese cabbage the type II family which has expanded.

The physical map positions of the MADS-box genes on Chinese cabbage chromosomes were identified (Fig. 5). Among the 160 *BrMADS* genes, two genes (*BrMADS150*, *BrMADS134*) could not be anchored on any of the Chinese cabbage chromosomes. *BrMADS150* and *BrMADS134* are on Scaffold 000343 and 000385, respectively. The other 158 members of the *BrMADS* genes were distributed non-randomly on 10 Chinese cabbage chromosomes (Fig. 5a). Chromosomes 2 and 9 contain the most MADS-box genes (15/16 %), whereas chromosome 8 contains the fewest (6 %) (Fig. 5b). We also found that some MADS-box genes cluster together in a region of the chromosome. For example, 16 genes clustered in the end of chromosome 2, and almost of the genes belong to *BrMIKC*^{C}. Type I and type II also show a differential distribution on Chinese cabbage chromosomes. The type I genes are distributed evenly across all ten chromosomes, whereas genes from type II are

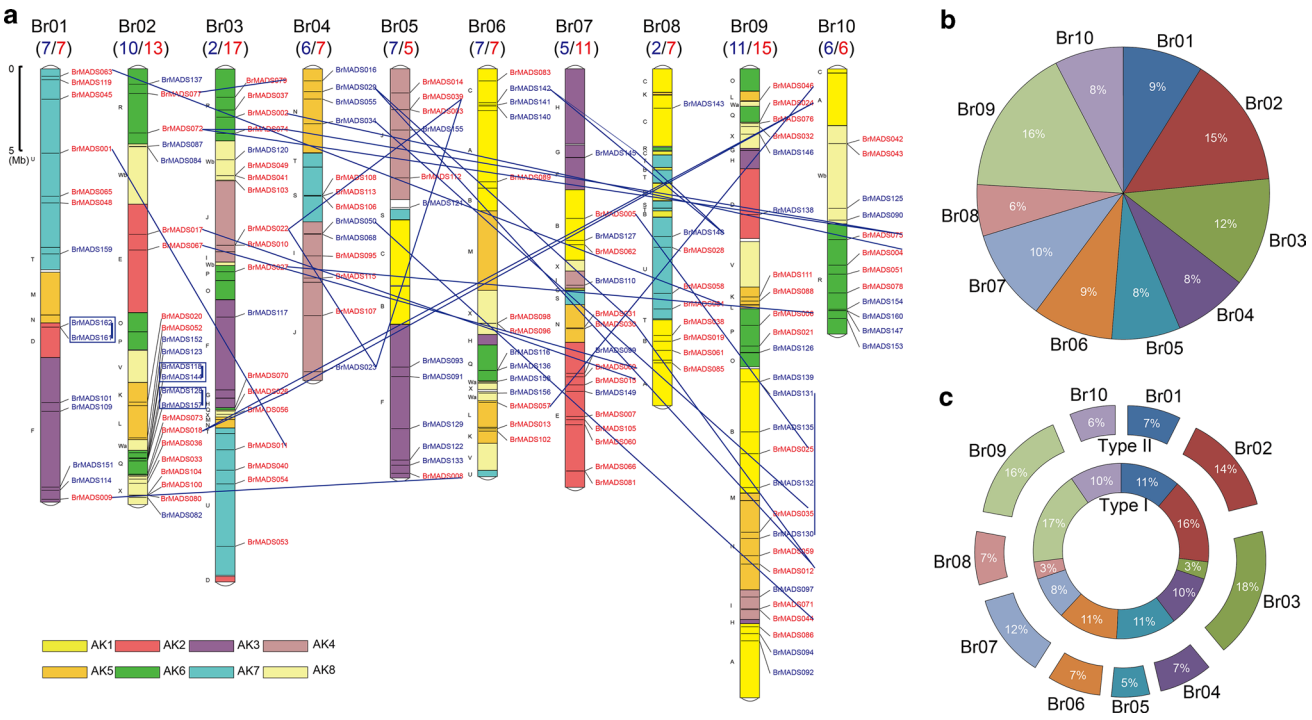

Fig. 5 Distribution of the *BrMADS* genes on ten Chinese cabbage chromosomes. **a** The 158 *BrMADS* genes non-randomly distributed on each conserved collinear blocks of the chromosome. Type I and type II genes are coloured *blue* and *red*, respectively Chromosome numbers are indicated above each chromosome followed by type I and type II numbers. The MADS-box genes present on duplicated chromosomal segments are connected by *blue lines* between the two relevant chromosomes. The tandem duplicated genes are in the *box*. The conserved collinear blocks on each chromosome are labeled *A–X* and are *colour-coded* according to inferred ancestral chromosomes following an established convention. **b** The percentages of *BrMADS* genes on each chromosome are demonstrated by the *pie*. **c** The percentages of *BrMADS* type I and type II genes on each chromosome are demonstrated by the *doughnut chart* (colour figure online)

located more members that from type I on chromosomes 03, 07 and 08 (Fig 5c). Interestingly, this is different from *Arabidopsis*, where the MIKC genes are distributed evenly across all five chromosomes, whereas the type I genes are located mainly on chromosomes 1 and 5 (Parenicová et al. 2003).

Duplicated genes from eukaryotic transcription factor families have originated predominantly from inter-chromosomal duplications (Friedman and Hughes 2001). The large size of the gene family MADS-box in *B. rapa* may suggest that this gene family underwent frequent duplication events during evolution. To learn more regarding the duplication of these genes, we defined the duplicated genes based on their *Ks* values and phylogenetic criteria (Supplementary Table 7, 8 and 9). Furthermore, these genes, which share similar gene structure and protein structure, were shown in chromosomes and in the phylogenetic tree (Fig. 5 and Supplementary Fig. 8). The duplicated genes were clustered closely together at the extremities of the phylogenetic tree. Most MADS-box genes have undergone segment duplication (39 duplications), whereas others have undergone tandem duplication (6 duplications) (Fig. 5). Tandem duplications have produced MADS-box gene clusters or hotspots, whereas segment duplications have produced many homologs of MADS-box genes on different chromosomes, as indicated with purple lines. The results indicated that the divergence time of duplicated *BrMADS* gene pairs ranged from 0.18 to 11.29 million years ago (MYA) and averaged 6.43 MYA, which indicates that the duplicated divergence of the MADS family members in *B. rapa* mostly accompanied the triplication events (5–9 MYA) (Supplementary Table 9) (Wang et al. 2011).

Differential expression of *BrMADS* genes in various tissues

To identify tissue-specific expression profiles of *BrMADS* genes, we utilised transcriptome data that were derived from Illumina RNA-Seq reads that were generated and analysed by Tong et al. (2013). The transcript abundance of 160 *BrMADSs* in 6 different tissues, including callus, root, stem, leaf, flower, and silique, was obtained; however, almost all of the type I *BrMADSs* either transcribed at too low a level to be detected or have spatial and temporal expression patterns that had no expression in the RNA-seq

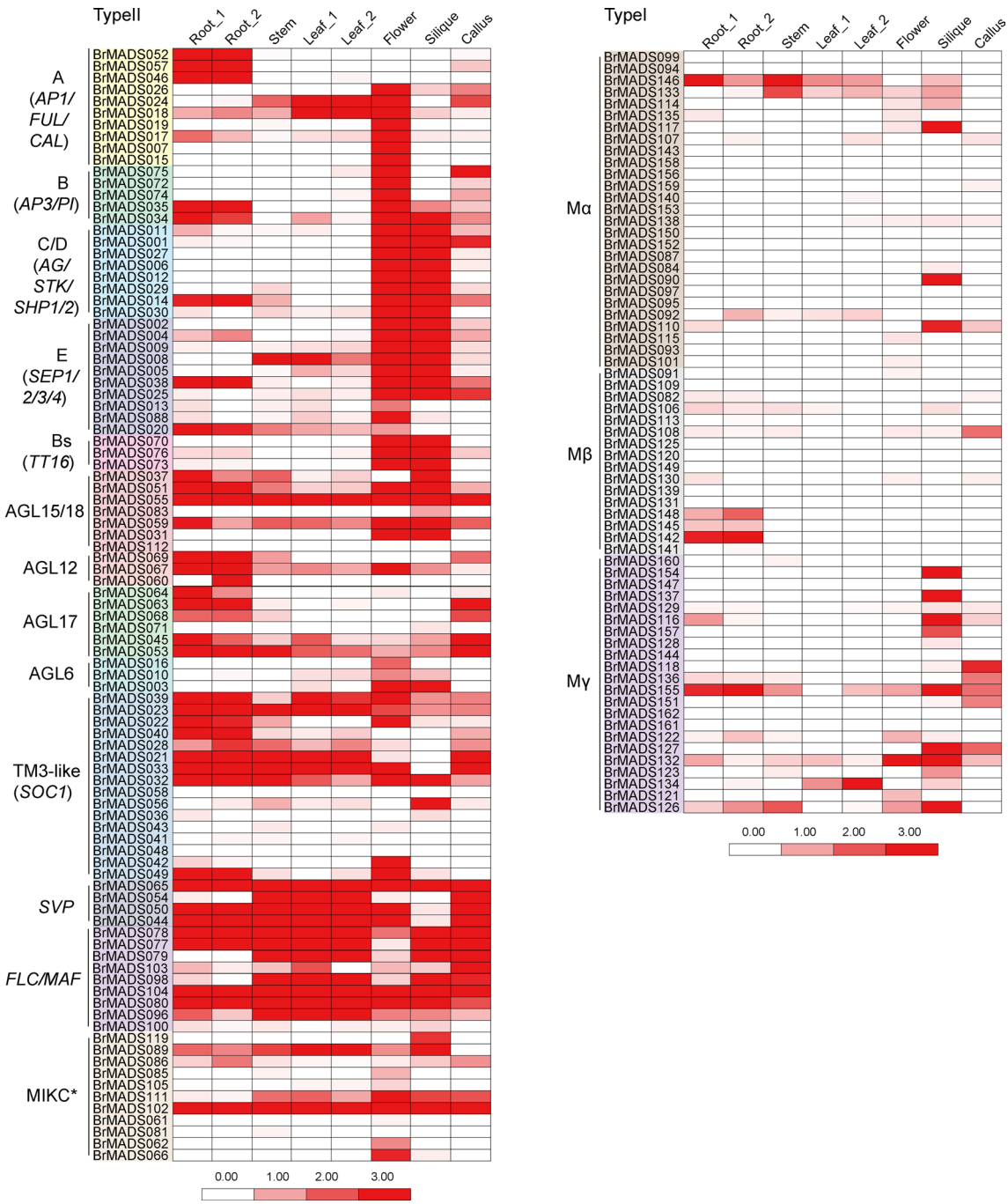

Fig. 6 Heat map representation of *BrMADS* genes in various tissues. The tissues included callus, root, stem, leaf, flower, and silique. Two samples of root and leaf tissues were generated from different batches of plants. The *bar* at the bottom of each heat map represents relative expression values

libraries (Supplementary Table 10). Of the *BrMIKC* genes, some exhibit tissue-specific expression (Fig. 6). The differences expression levels between typeI and typeII genes in the Chinese cabbage tissues were consistent with that in *Arabidopsis* (Kofuji et al. 2003). In Chinese cabbage the well-known ABCDE model genes have high expression levels in flowers, except *BrMADS046, 052, 057,* which only have high expression levels in roots. The CDE genes that are not only in flowers but also in siliques have higher expression levels. Some genes, which include *BrMADS077* and *078* (*FLC*), only had low expression in flowers, suggesting that these genes negatively regulate the flowering time similar to results of a previous report (Michaels and Amasino 1999). *BrMADS045, 053, 063, 064, 068* (*AGL17*)

have high expression levels in roots, which may be associated with root development (Zhang and Forde 1998). There are some genes that are expressed in all of the six tissues; these genes are *BrMADS023, 039, 032* (TM3-like); *BrMADS065*, (*SVP*); *BrMADS080, 096, 104* (*MAF*); and *BrMADS055* (*AGL18*). *SOC* and *SVP* control the flowering time, although *SOC* is a positive regulator, in contrast with *SVP*, *FLC* and *MAF*. *AGL18* is a floral transition gene. The expression patterns of these genes may be related to their roles in plant growth and development. The MIKC* group genes had low expression levels in these six tissues, except *BrMADS102*. Perhaps, this gene is also involved in the mediation of plant growth and development.

Differential expression of *BrMADS* genes under stresses

In the process of plant growth and development, plants are subject to various abiotic stresses. Thus, to identify the stress-responsive *BrMADS* genes under cold, heat, GA, SA, and ABA treatments in the seedling stage, we performed comprehensive expression profiling of *BrMIKC* genes using qRT-PCR (Supplementary Table 12), and the results are shown using the program TreeView. The expression levels of several genes were upregulated, whereas most genes were downregulated or showed no change under these treatments (Fig. 7). Under ABA treatment, some genes were upregulated at 4 h but were downregulated at

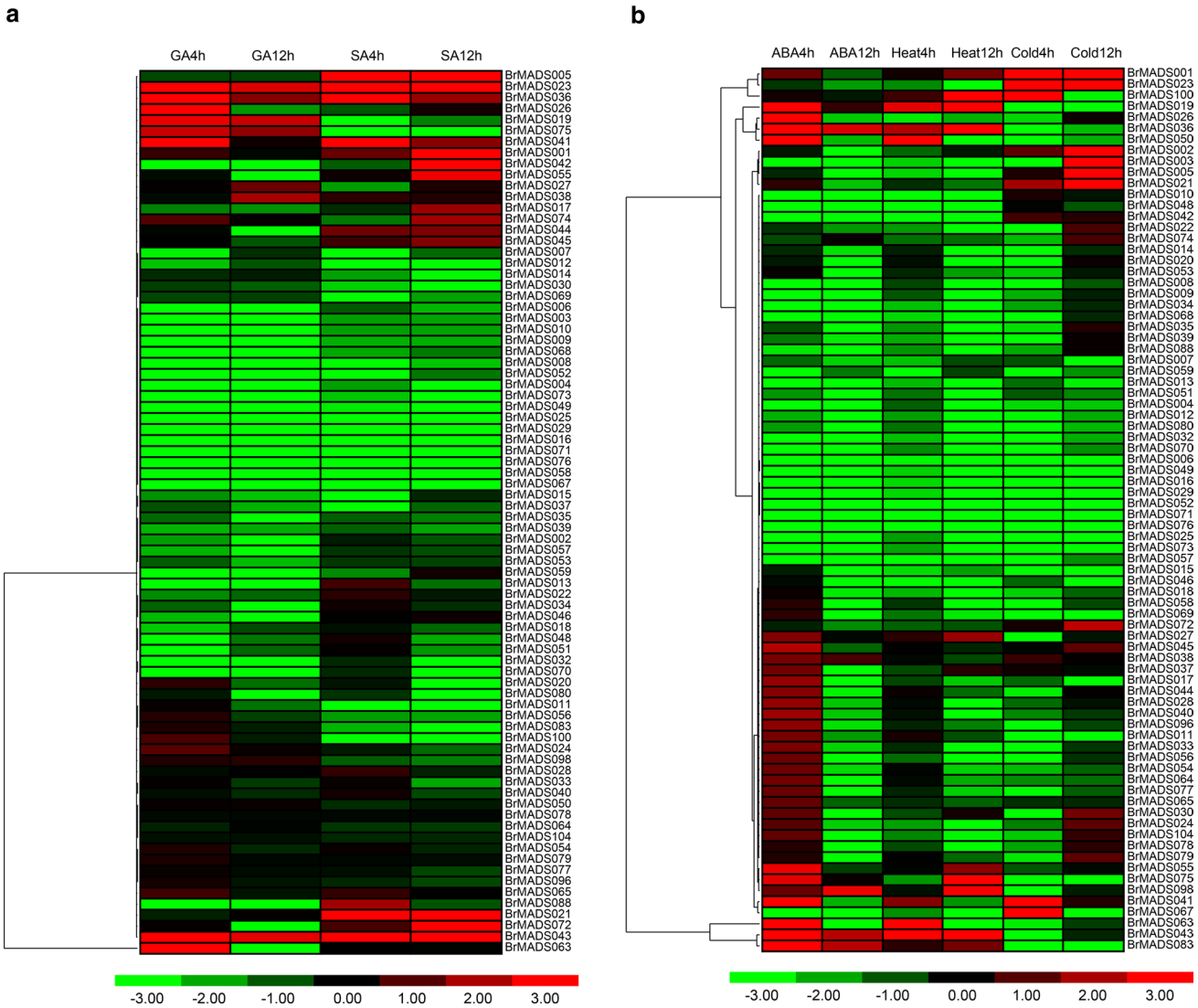

Fig. 7 Expression analysis of Chinese cabbage MIKC genes under five abiotic stresses. Heat map representation and hierarchical clustering of *BrMIKC* genes during **a** GA and SA stresses, **b** ABA, cold and heat stresses. Every stress contains two times, 4 and 12 h. The relative expression levels of *BrMIKCs* in leaves under these stresses were quantified against the control transcript levels. The *bar* at the *bottom* of each heat map represents relative expression values

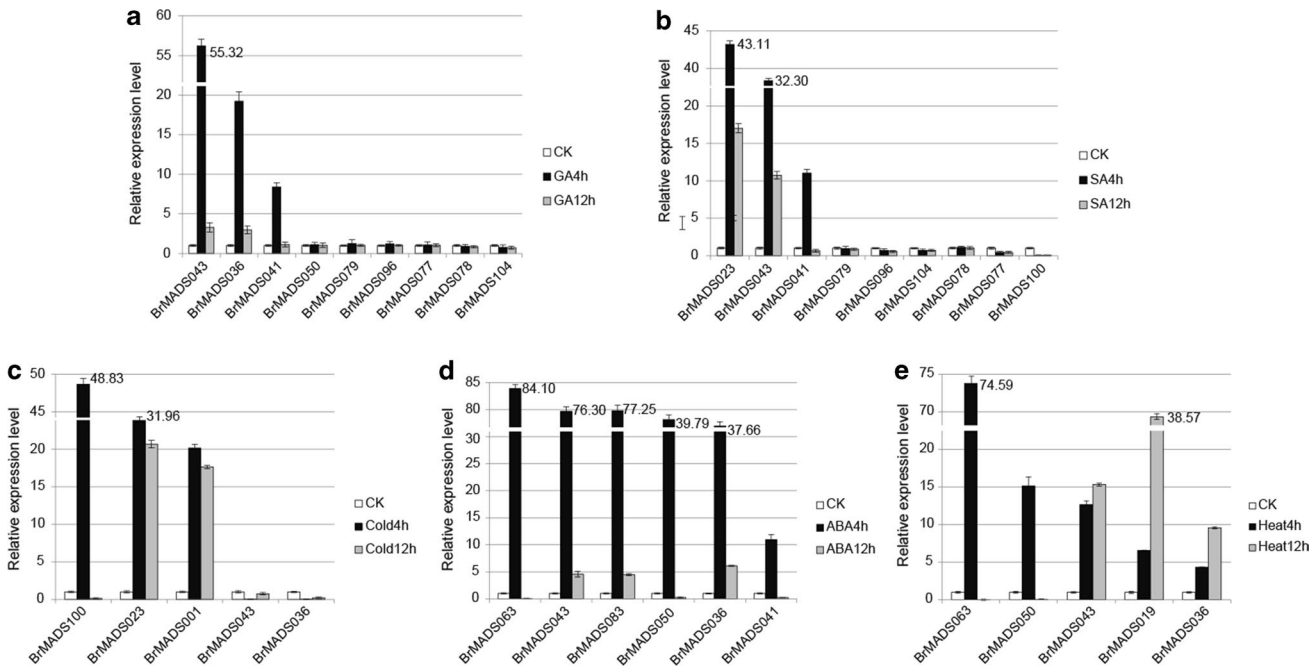

Fig. 8 The relative expression ratio of *BrMIKC* genes under abiotic stresses. **a** GA; **b** SA; **c** ABA; **d** cold; **e** heat. *Error bars* represent standard errors from three independent replicates

12 h. Because of the many *BrMIKC* genes and their complicated expression changes, only some were selected for further analysis (Fig. 8).

Three TM3-like ortholog *BrMADS* genes (*BrMADS043*, *BrMADS036* and *BrMADS041*) were upregulated at 4 and 12 h under GA treatment. The remaining six genes (*BrMADS050, SVP; 079, FLC; 096, MAF; 077, FLC; 078, FLC* and *104, MAF*) had no response to GA (Fig. 8a). Under SA treatment, *BrMADS023, 043,* and *041*, which are TM3-like ortholog genes, were upregulated at 4 and 12 h. *BrMADS079, FLC; 096, MAF; 104, MAF; 078, FLC; 077, FLC* and *100, MAF* were found with no response to SA treatment (Fig. 8b).

Under the other three abiotic stress treatments, most *BrMADS* genes have relatively low expression levels (Fig. 7b), whereas several genes were significantly upregulated in these different treatment conditions (Fig. 8c–e). Under cold stress, the *MAF* ortholog, *BrMADS* gene *BrMADS100*, was found with a high expression at 4 h but was downregulated at 12 h. The expression of two *BrMADS* genes (*BrMADS001, AG* and *023, SOC*) was over 20 times that of the control at 4 and 12 h. In contrast, two genes (*BrMADS043, SOC* and *036, SOC*) were downregulated at 4 and 12 h. Under ABA stress, the relative expression levels of six *BrMADS* genes (*BrMADS063, AGL17; BrMADS043, SOC; BrMADS083, AGL18; BrMADS050, SVP; BrMADS036, SOC* and *BrMADS041, SOC*) were upregulated at 4 h and then decreased rapidly at 12 h. Under heat stress, the expression levels of five *BrMADS* genes were upregulated; among these genes, *BrMADS063, AGL17* and *BrMADS050, SVP* had a higher expression levels at 4 h than at 12 h. Interestingly, *BrMADS043, SOC; BrMADS019, CAL* and *BrMADS036, SOC* had opposite results.

From the results above, 45 duplicated genes were found in the Chinese cabbage MADS-box gene family (Fig. 5 and Supplementary Fig. 8). Based on the expression levels of these genes (34 *BrMIKC* genes) under different stresses in Chinese cabbage, a line chart was used to show the different trends among them (Fig. 9 and Supplementary Table 13). Interestingly, we found four gene pairs with similar expression, including *BrMADS* (008:009; 034:035; 017:015; 077:079). Finally, the comparison of expression profiles of the duplicated gene pairs revealed that most duplicated genes had diverged significantly in expression level except for the above-mentioned four gene pairs.

Discussion

Chinese cabbage is one of the most important vegetables that are cultivated worldwide. For successful breeding of biennial plants like Chinese cabbage, flowering time and floral organ development are important traits. MADS-box genes are important regulators of both flowering and floral organ development, and are thus key factors for the

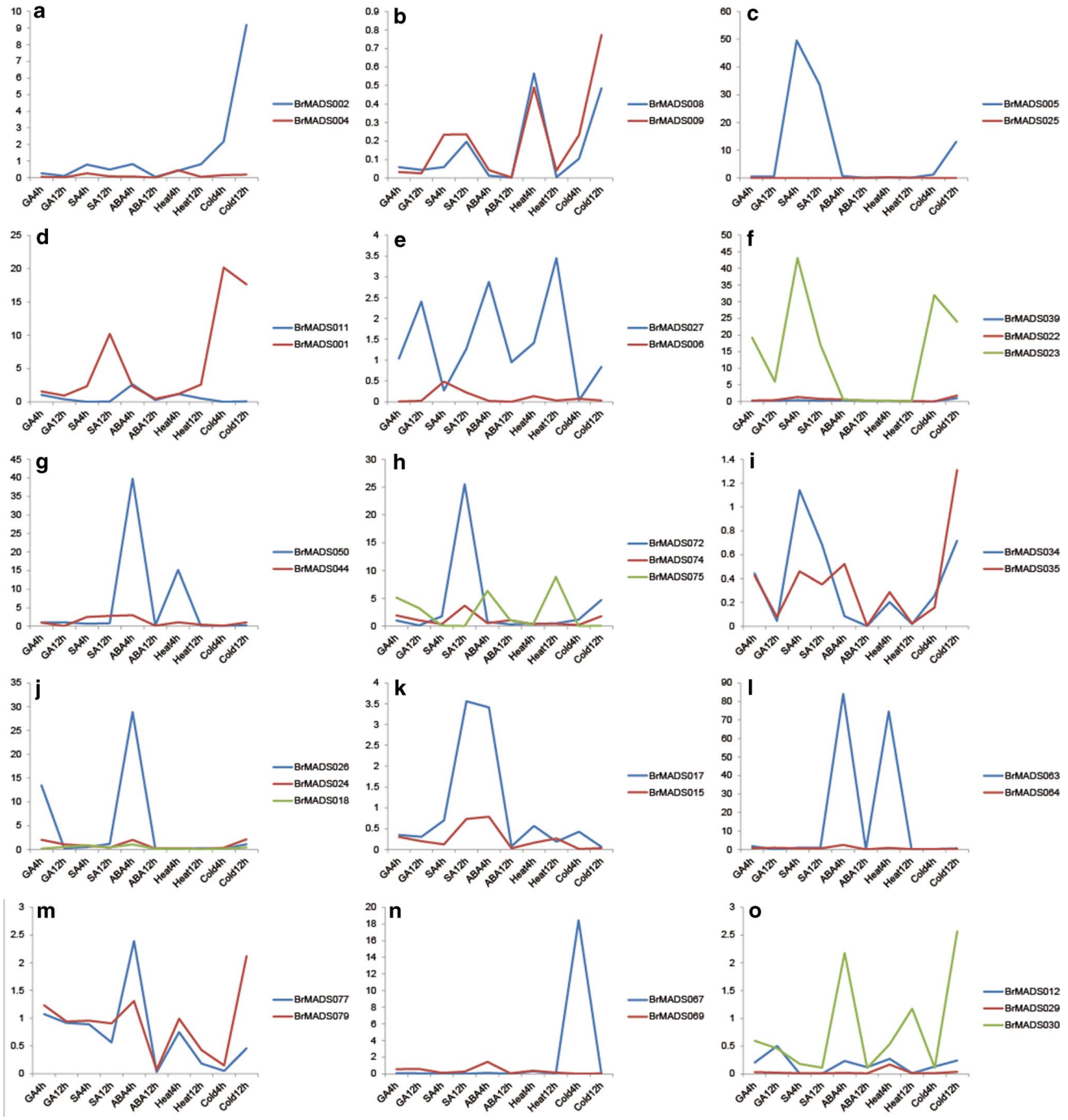

Fig. 9 MADS-box genes lying on duplicated segments of Chinese cabbage genome and their expression pattern. Expression patterns of duplicated genes have been compared in this figure. *X*-axis represents the abiotic stresses. *Y*-axis represents the raw expression value

breeding of Chinese cabbage. Based on sequence similarities and the complete genome sequence, attempts have been made to predict the function of MADS-box genes in diverse species (Wang et al. 2011). The completion of the *B. rapa* genome sequence also enabled us to predict a function for the *BrMADS* genes, based on results of phylogenetic analyses and their high sequence similarity with their *Arabidopsis* orthologs. In this study, we have presented a number of Chinese cabbage MADS-box gene phylogenetic relationships with *Arabidopsis* genes, a detailed analysis regarding the MADS-box genes that were collected among 22 species and the expression of *BrMADS* genes.

First, we identified 160 genes as Chinese cabbage MADS-box family genes. These genes were classified into

five subfamilies, including MIKC^{C}, MIKC*, Mα, Mβ and Mγ. Furthermore, the subfamily MIKC can also be divided into 13 subgroups, similar to that of *Arabidopsis* and rice MIKC genes. However, the number of each subgroup in Chinese cabbage is different compared with *Arabidopsis* and rice. This result implies that the retention of duplicates in the different clades has been differently in the species, such as *Arabidopsis*, rice, grapevine and Chinese cabbage. Thus, different groups of MADS-box genes with different gene numbers are under different evolutionary constraints (Nam et al. 2003; Airoldi and Davies 2012).

Most higher organisms pass through different ploidy levels at different development stages (Tang et al. 2008). Genome duplication has permanently shaped the architecture and function of many higher eukaryotic genomes. In the course of evolution, some duplicated genes have been retained, whereas others lost their function (Nakano et al. 2006). Generally, transcriptional regulators within the same group have recent common evolutionary origins and share conserved motifs with similar molecular functions (Doebley and Lukens 1998; Theissen et al. 2000). To understand the evolution of the MADS-box transcription factors family, 1,848 genes were identified and analysed among 22 species in our research. Through the evolution of species, we found that angiosperms contain a relatively large number of MADS-box genes. Thus, some important events during the phylogeny of species, particularly spermatophytes, can be shown through the evolution of the MADS gene family. During the phylogeny of species, after the Gamma whole genome triplication (WGT) event, when the monocots and eudicots diverged, the Brassicaceae genome experienced two whole genome duplication (WGD) events, *Arabidopsis* beta and *Arabidopsis* alpha (Bowers et al. 2003). Later, the Chinese cabbage genome experienced one WGT event against *Arabidopsis* (Wang et al. 2011). Thus, the MADS-box genes experienced WGD events and WGT events, which may be the reason for the different numbers of MADS genes among different species. These genes are the products of two different types of duplications: type II genes have been mainly duplicated through whole genome duplications; type I genes have been mainly duplicated by smaller scale and more recent duplications (Airoldi and Davies 2012). In this study only the number of type II MADS-box genes in Chinese cabbage is more than that of *Arabidopsis*, and the number of type I genes is similar. For *BrMADS* genes, we find that type II MADS-box genes are preferentially retained relative to type I genes, whereas most of type I genes were not located in the syntenic regions (Supplementary Table 4). Therefore, type II MADS-box genes are conserved and have accumulated in the Chinese cabbage genome probably by natural and artificial selection. The evolution of the MADS gene family is probably related to its important role in plant growth and development, such as flowering and floral organ development.

Finally, the expression patterns of *BrMIKC* genes under stresses were identified based on the qRT-PCR analysis. The MADS-box gene family functions in almost every developmental process in plants, and because plants are in the external environment exposed to various abiotic stresses, the MADS-box genes may be affected by some of these stresses. In previous reports, MADS-box genes have been shown to be affected by low temperature stress in tomato (Lozano et al. 1998) and by application of hormones like GA (Moon et al. 2003a, b), SA (Martínez et al. 2004) and ABA (Puig et al. 2013) in other plants. GA functions not only to promote the growth of plant organs but also to induce phase transitions during development. GA promotes flowering in *Arabidopsis* through the activation of genes that encode the floral integrator *Suppressor of Overexpression of Constans 1* (*SOC1*) in the inflorescence, in floral meristems, and in leaves (Mutasa-Göttgens and Hedden 2009). SA also regulates flowering in non-stressed plants because SA-deficient plants are late flowering. The regulation of the flowering time by SA seems to involve the photoperiod and autonomous pathways; however, SA regulation of flowering time does not require the function of the flowering time gene *Flowering Locus C* (*FLC*) (Martínez et al. 2004). In this study, we tested the stress-responsiveness of the *BrMADS* genes. We found that some *SOC1* ortholog genes were highly expressed during GA and SA treatments, while the *MAF* and *FLC* ortholog genes were not affected. *AGL17* may control plant vegetative development (Zhang and Forde 1998) and was induced by ABA treatment (Puig et al. 2013). We found that the *AGL17* ortholog gene *BrMADS063* was not only highly expressed under ABA (4 h) treatments but also highly expressed under heat (4 h) treatments. The results suggested that *BrMADS063* might be affected by heat regulation. The *MAF* ortholog gene *BrMADS100* was found to have high expression under cold (4 h) but was downregulated at 12 h. Therefore, we speculated that this gene might respond to cold stress. These results show that *BrMADS* genes may also function in the response to abiotic stresses, an interesting observation that has to be functionally explored further in the future.

In summary, 160 MADS-box transcription factors, including 65 type I and 95 type II genes, were identified in the entire Chinese cabbage genome. Type II MADS genes have been preferentially retained over type I genes, consistent with the gene dosage hypothesis. The isolation and identification of these transcription factors are likely to assist in clarifying the molecular genetic basis for Chinese cabbage genetic improvement and to provide functional gene resources for transgenic research. A comparison of phylogenetic relationships between Chinese cabbage and *Arabidopsis* MADS-box genes suggested that although most of the

basic subfamilies have been retained in Chinese cabbage, the number of each subfamily is different. The distributions and divergence of the duplicated gene pairs in *BrMADSs* supported the recent triplicated copies of the Chinese cabbage genome In addition, these analyses may also provide new opportunities to discover the tolerance mechanisms of the Chinese cabbage MIKC gene under stress conditions. The bioinformatics analysis results provides basic resources to examine the molecular regulation of Chinese cabbage development and stress resistance. The new information that has been generated is expected to help in the selection of appropriate candidate genes for further functional characterisation.

Acknowledgments This work is supported by the National Basic Research Program of China (973 Program, 2012CB113903), the Fundamental Research Funds for the Central Universities of China (KYZ201111), the National Natural Science Foundation of China (31272173, 1301782) and the Jiangsu Province Natural Science Foundation (BK20130673).

References

Adamczyk BJ, Lehti-Shiu MD, Fernandez DE (2007) The MADS domain factors AGL15 and AGL18 act redundantly as repressors of the floral transition in Arabidopsis. Plant J 50:1007–1019

Airoldi CA, Davies B (2012) Gene duplication and the evolution of plant MADS-box transcription factors. J Genet Genomics 39:157–165

Alvarez-Buylla ER, Liljegren SJ, Pelaz S, Gold SE, Burgeff C, Ditta GS, Vergara-Silva F, Yanofsky MF (2000) MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. Plant J 24:457–466

Arora R, Agarwal P, Ray S, Singh A, Singh V, Tyagi A, Kapoor S (2007) MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. BMC Genom 8:242

Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208

Becker A, Theißen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol Phylogenet Evol 29:464–489

Bemer M, Heijmans K, Airoldi C, Davies B, Angenent GC (2010) An atlas of type I MADS box gene expression during female gametophyte and seed development in *Arabidopsis*. Plant Physiol 154:287–300

Bezerra IC, Michaels SD, Schomburg FM, Amasino RM (2004) Lesions in the mRNA cap-binding gene ABA HYPERSENSITIVE 1 suppress FRIGIDA-mediated delayed flowering in *Arabidopsis*. Plant J 40:112–119

Birchler JA, Veitia RA (2007) The gene balance hypothesis: from classical genetics to modern genomics. Plant Cell 19:395–402

Bowers JE, Chapman BA, Rong J, Paterson AH (2003) Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422:433–438

Cheng F, Wu J, Fang L, Wang X (2012) Syntenic gene analysis between *Brassica rapa* and other Brassicaceae species. Front Plant Sci. 3:198

Cheng F, Mandáková T, Wu J, Xie Q, Lysak MA, Wang XW (2013) Deciphering the diploid ancestral genome of the mesohexaploid *Brassica rapa*. Plant Cell 25:1541–1554

Cho S, Jang S, Chae S, Chung KM, Moon YH, An G, Jang SK (1999) Analysis of the C-terminal region of *Arabidopsis thaliana* APETALA1 as a transcription activation domain. Plant Mol Biol 40:419–429

Colombo M, Masiero S, Vanzulli S, Lardelli P, Kater MM, Colombo L (2008) *AGL23*, a type I MADS-box gene that controls female gametophyte and embryo development in *Arabidopsis*. Plant J 54:1037–1048

De Bodt S, Raes J, Van de Peer Y, Theißen G (2003) And then there were many: MADS goes genomic. Trends Plant Sci 8:475–483

Díaz-Riquelme J, Lijavetzky D, Martínez-Zapater JM, Carmona MJ (2009) Genome-wide analysis of MIKCC-type MADS box genes in grapevine. Plant Physiol 149:354–369

Doebley J, Lukens L (1998) Transcriptional regulators and the evolution of plant form. Plant Cell 10:1075–1082

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863–14868

Finn RD, Clements J, Eddy SR (2011) HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–W37

Friedman R, Hughes AL (2001) Gene duplication and the structure of eukaryotic genomes. Genome Res 11:373–381

Heid CA, Stevens J, Livak KJ, Williams PM (1996) Real time quantitative PCR. Genome Res 6:986–994

Henschel K, Kofuji R, Hasebe M, Saedler H, Münster T, Theißen G (2002) Two ancient classes of MIKC-type MADS-box genes are present in the moss *Physcomitrella patens*. Mol Biol Evol 19:801–814

Hong JK, Kim SY, Kim KS, Kwon SJ, Kim JS, Kim JA, Lee SI, Lee YH (2012) Overexpression of a *Brassica rapa* MADS-box gene, *BrAGL20*, induces early flowering time phenotypes in *Brassica napus*. Plant Biotechnol Rep 7:231–237

Honma T, Goto K (2001) Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature 409:525–529

Jensen R, Bassham J (1966) Photosynthesis by isolated chloroplasts. Proc Natl Acad Sci USA 56:1095

Kofuji R, Sumikawa N, Yamasaki M, Kondo K, Ueda K, Ito M, Hasebe M (2003) Evolution and divergence of the MADS-box gene family based on genome-wide expression analyses. Mol Biol Evol 20:1963–1977

Köhler C, Hennig L, Spillane C, Pien S, Gruissem W, Grossniklaus U (2003) The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene *PHERES1*. Gene Dev 17:1540–1553

Krzywinski M, Schein J, Birol İ, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645

Lee JH, Yoo SJ, Park SH, Hwang I, Lee JS, Ahn JH (2007) Role of *SVP* in the control of flowering time by ambient temperature in *Arabidopsis*. Gene Dev 21:397–402

Lee TH, Tang H, Wang X, Paterson AH (2013) PGDD: a database of gene and genome duplication in plants. Nucleic Acids Res 41:D1152–D1158

Letunic I, Doerks T, Bork P (2012) SMART 7: recent updates to the protein domain annotation resource. Nucleic Acids Res 40:D302–D305

Li L, Stoeckert CJ, Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res 13:2178–2189

Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF (2000) SHATTERPROOF MADS-box genes control seed dispersal in *Arabidopsis*. Nature 404:766–770

Liu C, Chen H, Er HL, Soo HM, Kumar PP, Han JH, Liou YC, Yu H (2008) Direct interaction of *AGL24* and *SOC1* integrates flowering signals in *Arabidopsis*. Development 135:1481–1491

Liu T, Li Y, Ren J, Zhang C, Kong M, Song X, Zhou J, Hou X (2013) Over-expression of *BcFLC1* from non-heading Chinese

cabbage enhances cold tolerance in *Arabidopsis*. Biol Plantarum 57:262–266

Lozano R, Angosto T, Gomez P, Payan C, Capel J, Huijser P, Salinas J, Martinez-Zapater JM (1998) Tomato flower abnormalities induced by low temperatures are associated with changes of expression of MADS-Box genes. Plant Physiol 117:91–100

Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science 290:1151–1155

Martínez C, Pons E, Prats G, León J (2004) Salicylic acid regulates flowering time and links defence responses and reproductive development. Plant J 37:209–217

Masiero S, Colombo L, Grini PE, Schnittger A, Kater MM (2011) The emerging importance of type I MADS box transcription factors for plant reproduction. Plant Cell 23:865–872

Michaels SD, Amasino RM (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. Plant Cell 11:949–956

Moon J, Suh SS, Lee H, Choi KR, Hong CB, Paek NC, Kim SG, Lee I (2003a) The *SOC1* MADS-box gene integrates vernalization and gibberellin signals for flowering in *Arabidopsis*. Plant J 35:613–623

Moon J, Suh SS, Lee H, Choi KR, Hong CB, Paek NC, Kim SG, Lee I (2003b) The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. Plant J 35:613–623

Mutasa-Göttgens E, Hedden P (2009) Gibberellin as a factor in floral regulatory networks. J Exp Bot 60:1979–1989

Nakano T, Suzuki K, Fujimura T, Shinshi H (2006) Genome-wide analysis of the ERF gene family in *Arabidopsis* and rice. Plant Physiol 140:411–432

Nam J, Ma H, Nei M (2003) Antiquity and evolution of the MADS-box gene family controlling flower development in plants. Mol Biol Evol 20:1435–1447

Nesi N, Debeaujon I, Jond C, Stewart AJ, Jenkins GI, Caboche M, Lepiniec L (2002) The TRANSPARENT TESTA16 locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. Plant Cell 14:2463–2479

Norman C, Runswick M, Pollock R, Treisman R (1988) Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response element. Cell 55:989–1003

Parenicová L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B (2003) Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis new openings to the MADS world. Plant Cell 15:1538–1551

Passmore S, Maine GT, Elble R, Christ C, Tye BK (1988) Saccharomyces cerevisiae protein involved in plasmid maintenance is necessary for mating of MATα cells. J Mol Biol 204:593–606

Pinyopich A, Ditta GS, Savidge B, Liljegren SJ, Baumann E, Wisman E, Yanofsky MF (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature 424:85–88

Puig J, Meynard D, Khong GN, Pauluzzi G, Guiderdoni E, Gantet P (2013) Analysis of the expression of the AGL17-like clade of MADS-box transcription factors in rice. Gene Expr Patterns 13:160–170

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C, Pang N, Forslund K, Ceric G, Clements J (2012) The Pfam protein families database. Nucleic Acids Res 40:D290–D301

Pylatuik JD, Lindsay DL, Davis AR, Bonham-Smith PC (2003) Isolation and characterization of a *Brassica napus* cDNA corresponding to a B-class floral development gene. J Exp Bot 54:2385–2387

Ratcliffe OJ, Kumimoto RW, Wong BJ, Riechmann JL (2003) Analysis of the *Arabidopsis* MADS AFFECTING FLOWERING gene family: *MAF2* prevents vernalization by short periods of cold. Plant Cell 15:1159–1169

Riechmann JL, Meyerowitz EM (1997) MADS domain proteins in plant development. Biol Chem 378:1079–1102

Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G (2000) Distinct roles of CONSTANS target genes in reproductive development of *Arabidopsis*. Science 288:1613–1616

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13:2498–2504

Sommer H, Beltran JP, Huijser P, Pape H, Lönnig W, Saedler H, Schwarz-Sommer Z (1990) Deficiens, a homeotic gene involved in the control of flower morphogenesis in Antirrhinum majus: the protein shows homology to transcription factors. EMBO J 9:605

Song XM, Huang ZN, Duan WK, Ren J, Liu TK, Li Y, Hou XL (2013) Genome-wide analysis of the bHLH transcription factor family in Chinese cabbage (*Brassica rapa* ssp. *pekinensis*). Mol Genet Genomics 289:77–91

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Tang H, Bowers JE, Wang X, Ming R, Alam M, Paterson AH (2008) Synteny and collinearity in plant genomes. Science 320:486–488

Tapia-López R, García-Ponce B, Dubrovsky JG, Garay-Arroyo A, Pérez-Ruíz RV, Kim SH, Acevedo F, Pelaz S, Alvarez-Buylla ER (2008) An AGAMOUS-related MADS-box gene, XAL1 (*AGL12*), regulates root meristem cell proliferation and flowering transition in Arabidopsis. Plant Physiol 146:1182–1192

Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Münster T, Winter KU, Saedler H (2000) A short history of MADS-box genes in plants. Plant Mol Biol 42:115–149

Thomas BC, Pedersen B, Freeling M (2006) Following tetraploidy in an Arabidopsis ancestor, genes were removed preferentially from one homeolog leaving clusters enriched in dose-sensitive genes. Genome Res 16:934–946

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Tong C, Wang X, Yu J, Wu J, Li W, Huang J, Dong C, Hua W, Liu S (2013) Comprehensive analysis of RNA-seq data reveals the complexity of the transcriptome in *Brassica rapa*. BMC Genom 14:689

Wang X, Wang H, Wang J, Sun R, Wu J, Liu S, Bai Y, Mun JH, Bancroft I, Cheng F (2011) The genome of the mesopolyploid crop species *Brassica rapa*. Nat Genet 43:1035–1039

Wilmowicz E, Kęsy J, Kopcewicz J (2008) Ethylene and ABA interactions in the regulation of flower induction in Pharbitis nil. J Plant Physiol 165:1917–1928

Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM (1990) The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. Nature 346:35–39

Zahn LM, Feng B, Ma H (2006) Beyond the ABC-model: regulation of Floral Homeotic Gene. Adv Bot Res 44:163–207

Zhang H, Forde BG (1998) An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. Science 279:407–409

Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J (2006) KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinform 4:259–263