

RESEARCH ARTICLE

Open Access



# Genome-wide identification of MADS-box family genes in moso bamboo (*Phyllostachys edulis*) and a functional analysis of *PeMADS5* in flowering

Yuting Zhang<sup>1</sup>, Dingqin Tang<sup>1</sup>, Xinchun Lin<sup>1</sup>, Mingquan Ding<sup>2\*</sup> and Zaikang Tong<sup>1\*</sup>

## Abstract

**Background:** MADS-box genes encode a large family of transcription factors that play significant roles in plant growth and development. Bamboo is an important non-timber forest product worldwide, but previous studies on the moso bamboo (*Phyllostachys edulis*) MADS-box gene family were not accurate nor sufficiently detailed.

**Results:** Here, a complete genome-wide identification and characterization of the MADS-box genes in moso bamboo was conducted. There was an unusual lack of type-I MADS-box genes in the bamboo genome database (<http://202.127.18.221/bamboo/index.php>), and some of the *PeMADS* sequences are fragmented and/or inaccurate. We performed several bioinformatics techniques to obtain more precise sequences using transcriptome assembly. In total, 42 MADS-box genes, including six new type-I MADS-box genes, were identified in bamboo, and their structures, phylogenetic relationships, predicted conserved motifs and promoter *cis*-elements were systematically investigated. An expression analysis of the bamboo MADS-box genes in floral organs and leaves revealed that several key members are involved in bamboo inflorescence development, like their orthologous genes in *Oryza*. The ectopic overexpression of one MADS-box gene, *PeMADS5*, in *Arabidopsis* triggered an earlier flowering time and the development of an aberrant flower phenotype, suggesting that *PeMADS5* acts as a floral activator and is involved in bamboo flowering.

**Conclusion:** We produced the most comprehensive information on MADS-box genes in moso bamboo. Additionally, a critical *PeMADS* gene (*PeMADS5*) responsible for the transition from vegetative to reproductive growth was identified and shown to be related to bamboo floral development.

**Keywords:** MADS-box, Bamboo, Gene expression pattern, Flower time, Floral organ

## Background

MADS-box genes encode a large family of transcription factors that have essential roles in animals, plants, and fungi [1]. The first plant MADS-box genes were found to regulate floral meristem identity. Since then, several MADS-box genes have been reported to control the vegetative to reproductive phase transition in plants and developmental processes in plant organs, such as fruit, root, stem, and leaf [2–4]. This large gene family

contains two groups, type I (SRF-like) and type II (MEF2-like), based on their conserved domains [5]. In plants, the type-I genes encode plant-specific transcription factors of the M $\alpha$ , M $\beta$ , and M $\gamma$  subfamilies, while the most well-known MADS genes are from the type-II group, such as the MIKC<sup>C</sup>-type and MIKC<sup>\*</sup>-type [6, 7]. The term MIKC originated from the four major domains of proteins encoded by type-II genes: the MADS (M) domain followed by an Intervening (I) domain, a second most conserved Keratin (K) domain and a C-terminal domain [8].

Flowering is a complicated process that requires co-operation and interaction among numerous genes. Loss-of-function plant mutations have been used to identify

\* Correspondence: 20110028@zafu.edu.cn; zktong@zafu.edu.cn

<sup>2</sup>The Key Laboratory for Quality Improvement of Agricultural Products of Zhejiang Province, College of Agriculture and Food Science, Zhejiang A & F University, Lin'an, Zhejiang, China

<sup>1</sup>State Key Laboratory of Subtropical Silviculture, Zhejiang A & F University, Lin'an, Zhejiang, China



many MADS-box genes that play crucial roles in flowering processes, including flowering time (SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC), FLOWERING LOCUS C (FLC), and SHORT VEGETATIVE PHASE (SVP)) [9–11], formation of floral meristems and organs (*APETALA1* (AP), *FRUITFULL* (FUL), *PISTILLATA* (PI), *APETALA3* (AP<sub>3</sub>), *SHATTERPROOF* (SHP) 1/2, *SEPALLATA* (SEP) 1/2/3 and *AGAMOUS* (AG)) [12–16], fruit ripening (SHP1 and SHP2) [17], seed pigmentation, and embryo development (TRANSPARENT TESTA16) [8, 18]. The ABCDE model explains the ontogeny of dicot flowers [16, 19]. In general, A-class genes participate in sepal development, whereas C-class genes are involved in carpel development. Petal development requires both A- and C-class genes, whereas stamen development needs both B- and C-class genes. D-lineage genes are necessary for ovule development [15, 20, 21]. E-class genes redundantly specify all floral organ identities and floral meristem determination [16]. Almost every gene from this model, except *Arabidopsis* AP<sub>2</sub>, belongs to the type-II MADS-box subfamily, indicating that MADS-box genes play pivotal roles during flowering. Compared with type-II MADS-box genes, type-I genes in plants have attracted less attention [22]. However, functional studies of type-I MADS-box genes have revealed their crucial roles during plant reproduction and development, especially in determining female gametophyte, embryo, and endosperm development in *Arabidopsis* [22].

Bamboo is an important non-timber forest product worldwide. The plant is beneficial to the human food supply and several industries, as well as to the conservation of the environment and animal habitats. Moso bamboo (*Phyllostachys edulis*), native to China, is a large woody bamboo species with a complex rhizome system, hollow and highly lignified culms, and bisexual spikelets [23]. For the many years of the long vegetative phase, moso bamboo culms (building materials) and young shoots (edible vegetable) can be continually harvested. However, moso bamboo only flowers once and dies after seed production (monocarpy) [24]. Woody bamboos typically exhibit synchronous flowering in a community in which all culms flower and die within the same year [25]. The switch from vegetative to reproductive growth is difficult to predict. These unique characteristics have led to centuries of study by scientists and enthusiasts. Several hypotheses have been proposed for the mechanism of bamboo flowering, and physiological and biochemical studies have been conducted [23]. External controls, such as photoperiod and drought cycles, may initiate flowering [26, 27], and endogenous factors may control flowering induction, including the circadian clock, which has been supported by observations of synchronous flowering in parental stocks and transplanted

seedlings [28, 29]. However, the mechanism underlying the molecular regulation of flowering remains unclear.

In eudicots, the expression of some MADS-box genes can affect flowering time and the development of floral structures, which make MADS-box genes good candidates to understand the unique flowering patterns in bamboos. Lin identified two AP/SQUA-like MADS-box genes from *Phyllostachys praecox*, *PpMADS1* (FUL3 subfamily) and *PpMADS2* (FUL1 subfamily) [30]. Both genes play vital roles in the floral transition of bamboo. Next-generation sequencing allows the investigation of functional genes and flowering pathways on a genome scale. In the caespitose bamboo *Bambusa edulis*, two sequencing platforms have been used and 16 MADS-box genes (*BeMADS*) identified [31]. Most *BeMADS* genes are highly expressed in floral organs and share similar expression patterns with their homologs in *Oryza sativa*. A transcriptome analysis in *P. edulis* identified 38 putative MADS-box transcription factors in panicle tissues [32]. Although the first version of the bamboo genome sequence is available, the annotations of several genes are still tentative owing to the limitations of the previously used sequencing and assembling technologies. Here, we report a novel and systematic study of the MADS-box gene family involving not only database sequence identification but also sequence complementation and correction. We identified 42 MADS-box genes and then investigated the structures, phylogenetic relationships, conserved protein motifs, duplications, and expression patterns of these genes during floral development in bamboo. The overexpression of an *AGL24* homologous gene, *PeMADS5*, in *Arabidopsis* resulted in early flowering and floral abnormalities. Yeast two-hybrid analyses revealed that *PeMADS5* can interact with *PeAP1* (*PeMADS2*) and *PeSOC1* (*PeMADS34*), which suggests that *PeAGL24* and *PeSOC1* act as dimers to promote flowering. This study will serve as a useful reference for further functional analyses of candidate genes involved in bamboo floral development.

## Methods

### Identification and classification of MADS-box genes

We downloaded *P. edulis* genome sequences from the Bamboo Genome Database (<http://202.127.18.221/bamboo/index.php>). MADS-box protein sequences from *Arabidopsis*, *Oryza*, and *Brachypodium* were obtained from published studies [8, 18, 33]. To acquire the maximum number of MADS-box domain containing sequences in moso bamboo, we built three different Hidden Markov Models (HMM) profiles to search the *P. edulis* protein dataset based on *Arabidopsis*, *Oryza*, and *Brachypodium*, respectively. Gene name query was also employed on NCBI and BambooGDS. All of these candidate MADS-box genes were checked manually to

remove the incomplete and redundant sequences. The nomenclature of putative type II bamboo MADS-box genes was assigned based on their scaffolds rankings.

#### Determination of presence and absence of type I MADS-box gene clades

In the first round of identification, type I MADS-box subfamily genes escaped detection in the HMM search. Therefore, the coding sequences (CDS) of Type I MADS-box from *Oryza* and *Brachypodium* were used as the probe sequences to blast against the bamboo genome as local databases. We obtained the best three BLAST results to check whether they represent MADS-box genes that were already in our dataset. If not, the CDS and protein sequences were predicted by Fgenesh + software [34] based on the corresponding genomic scaffold.

#### Correcting the incomplete *PeMADSs* sequences

The *P. edulis* transcriptome reads were downloaded from the NCBI SRA database (Accession: SRR4450542, SRR4450543, SRR4450544, SRR4450545, SRR4450546, SRR4450547, SRR4450548, SRR4450549, SRR4450550, SRR4450551). One of the methods to correct *PeMADSs* sequences was using Trinity software. After trimming low-quality sequences, transcriptome data was de novo assembled with the Trinity software using default parameters [35]. Because we are only interested in the *PeMADSs*, all of the CDS sequences of the MADS-box family genes were used as the references. Then the software of the Tablet was employed to exhibit the alignment results [36]. To identify the incomplete 5' and 3' end sequence of the genes, another method to complete the sequences was using in-house scripts to conduct e-Genome-walking process. To validate these complementary sequences, full-length CDS sequences of 16 *PeMADSs* were amplified (Primers in Additional file 1: Table S1) and cloned into pMD18-T vectors. After sequencing, multiple sequence alignments among the cloned, assembled and original sequences were conducted to explore the accuracy of the methods we used here.

#### Subcellular localization analysis

The coding sequences of *PeMADS23* and PH01002755G0230 without the stop codon were amplified and then subcloned into the p2GWY7 vector and fused in-frame with the Yellow Fluorescent Protein (YFP) sequence under the control of the *CaMV35S* promoter. The fusion constructs were introduced into *Arabidopsis* protoplasts prepared from 4-day suspension cells by using 40% polyethylene glycol (PEG) as described previously [37]. YFP fluorescence was observed with a laser scan confocal microscope. The transient expression assay was repeated three times.

#### Phylogenetic analysis

*P. edulis* MADS-box proteins were aligned with *Oryza*, *Arabidopsis* by MAFFT [38]. The alignment was cropped to remove low conserved regions and finally contained the M- and part of the K domain. An unrooted neighbor-joining (NJ) tree was constructed using the MEGA7 package [33, 39]. The tree nodes were evaluated by bootstrap analysis for 1000 replicates. Branches with less than 50% bootstrap values were collapsed.

#### Conserved motif analysis

The conserved motifs were investigated by MEME version 2.2 online tool (Multiple Expectation Maximization for Motif Elicitation) (<http://meme-suite.org/>) [40]. The parameters are set as follows: number of repetitions: any, the maximum number of motifs: 10, optimum motif width set to  $\geq 6$  and  $\leq 200$ . The motifs obtained were annotated using the Simple Modular Architecture Research Tool (SMART) and NCBI CD search program [41, 42].

#### Cis-element enrichment analysis

PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) was used to predict cis-regulatory elements in the *PeMADS* promoter (1.5 kb upstream from the translational start codon) and intron region [43]. In this study, we selected cis-element associated with core promoter elements, protein binding sites, hormones responses, tissue-specific elements, light responsive elements, abiotic and biotic stress responses, circadian responses and cell cycle regulation elements.

#### Collection of plant material

The inflorescence samples of flowering moso bamboo and the leaf samples from non-flowering plants were collected in Guilin, Guangxi province from June to July 2016. The identification of four stages of bamboo inflorescence development is based on the numbers of florets and anatomical structures (stages F1–4): the first floral bud formation, initial stage of inflorescence development (3–5 florets), maturation of inflorescence (~10 florets) and anthesis. The leaf tissues were collected from the non-flowering moso bamboo, under the same growth environment with the flowering plants. The panicle and leaf samples were frozen in liquid nitrogen immediately and stored at  $-70^{\circ}\text{C}$ . Total RNA was extracted using the RNAPrep pure Plant Kit (TianGen, China). Contaminating DNA was removed using DNase (TaKaRa Bio Inc., Japan). After checking the quality of purified RNA on an agarose gel, the RNA samples were quantified by NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., USA). RNA was reverse transcribed from 5  $\mu\text{g}$  of total RNA in 100  $\mu\text{L}$  of reaction volume using the PrimeScript™ RT reagent kit (TaKaRa Bio Inc.,

Japan) according to the manufacturer's instructions. The resulting cDNA was used for further experiments.

### Expression analysis of MADS-box genes

The primers used for MADS-box expression analysis were designed by Primer Premier 5 (Additional file 1: Table S2). The expression levels of target genes were detected with SYBR<sup>®</sup> Premix Ex Taq II (TaKaRa Bio Inc., Japan). All qPCR assays were carried out in a CFX-96-well Real-Time System (BioRad, USA). The reaction mixture consisted of 10  $\mu$ L SYBR Green mix, 200 nM of each primer, and 2  $\mu$ L of diluted cDNA in a final volume of 20  $\mu$ L. The qPCR protocol consisted of an initial thermal cycling step of 95  $^{\circ}$ C for 3 mins, followed by 40 cycles of denaturation at 95  $^{\circ}$ C for 10 s and annealing with a temperature gradient from 50 to 60  $^{\circ}$ C for 30 s. All experimental samples were repeated in triplicate. To normalize the variance among different samples, NTB was used as a housekeeping gene for data normalization [44, 45]. The raw Cycle threshold (Ct) values were calculated automatically by the Bio-Rad CFX Manager (version 2.3) using the  $2^{-\Delta\Delta ct}$  method [46].

### Construction of *PeMADS5* ectopic expression transgenic *Arabidopsis* lines

The full open reading frame of the *PeMADS5* was amplified from cDNA using the primer pairs *PeMADS5* (Additional file 1: Table S1) and cloned into pMD18-T Vector. The pCXSXN (FJ905214) vector containing the CaMV35S promoter and the *Nos* terminator was digested with *Xcm*I (New England Biolabs) [47]. The PCR product was amplified from pMD18-T followed by an A-addition procedure. The digested pCXSXN plasmid and the PCR-amplified product were ligated using T4 DNA ligases (NEB). All recombinant plasmids identified from each individual *E. coli* colone were verified by sequencing. This ectopic expression construct was named 35S::*PeMADS5*. Healthy wild-type *Arabidopsis* (Columbia-0) plants were grown on soil under long-day photoperiod conditions (16 h of light/8 h of darkness). The construct was transformed to *Agrobacterium tumefaciens* strain GV3101 and used to transform *Arabidopsis* by the floral dip method [48]. *Arabidopsis* seeds obtained after transformation were plated on one-half-strength Murashige and Skoog medium containing 30 mg hygromycin for selection. Rosette leaves of Col-0, 35S::*PeMADS5* #2 and #10 were obtained when plants bore a 1-cm-long inflorescence. The inflorescences from col-0 and transgenic plants were also collected on full flowering stage. Plant materials were sampled for qRT-PCR. Statistical analysis for the results from qPCR expression data and bolting days of *Arabidopsis* were carried out using a Student's *t*-test. All data are presented as mean  $\pm$  SD.  $p \leq 0.05$  and  $P \leq 0.01$  were considered statistically significant compared with the Columbia-0 (Col-0) wild-type.

### Yeast two-hybrid assay

The experimental procedures of yeast two-hybrid assay were performed using the Matchmaker Two-Hybrid System (Clontech). The full coding sequence of *PeMADS2*, 5, 16, 20 and 34 were cloned into both pGBKT7 and pGADT7 vectors. The resulting recombinant plasmids were introduced into yeast strains Y2HGold and Y187, respectively. Two-hybrid interactions were assayed on selective SD/-Trp/-Leu double-dropout (DDO). Potential interactions were tested on selectiveSD/-Trp/-Leu/-His/-Ade/X- $\alpha$ -gal (40 mg/ml) media supplemented with 5 mM 3-amino-triazole (3-AT).

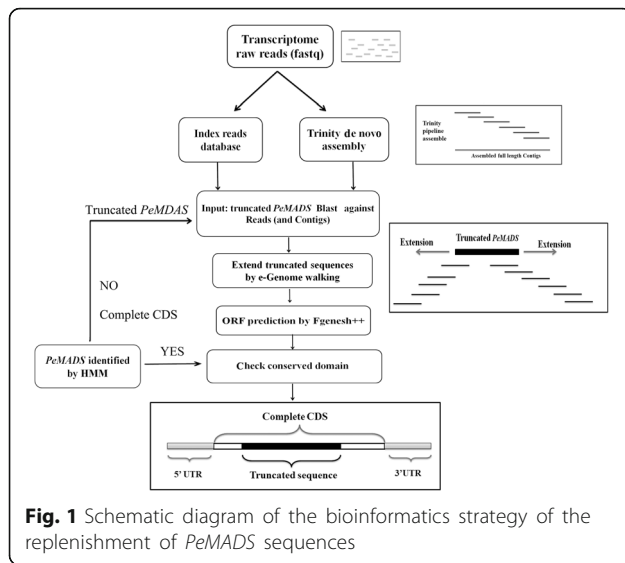
## Results

### Identification and replenishment of *P. edulis* MADS-box genes

The MADS-box protein sequences of three species, *Oryza*, *Arabidopsis*, and *Brachypodium*, were used to build a hidden Markov model, which was used to search the moso bamboo protein dataset. In total, 36 unique MADS-box proteins were identified (Table 1). *PeMADSs* (*PeMADS18*, -27, -30, -36, -37, -38, -42, and -44) do not contain the most conserved MADS domain, and they were verified to be pseudogenes. Furthermore, five type-I genes, a PI-like clade gene (*PeAP3*) and an AGL6-like clade gene (*PeAGL6*) were obtained by running BLAST algorithm-based searches of the Bamboo genome databases with *Oryza* and *Arabidopsis* homologous genes. In total, 42 *MADS* genes have been identified in *P. edulis*. However, owing to the limited accuracy of *Phyllostachys* scaffolding, some *PeMADSs* may lack full-length gene sequence information. Thus, we developed a bioinformatics pipeline to recover the full-gene regions using the transcriptome sequencing data. First, the transcriptome reads were indexed as the references, and then, the incomplete *PeMADS* sequences were used as query in BLAST algorithm-based searches of the reads database. The optimally aligned reads were chosen to extend the open reading frame of each gene. The newly modified sequences were then included in a new search cycle until no new reads could be identified for any extension. All of the replenished coding DNA sequences (CDSs) were translated by Fgenesh++ software (Fig. 1) [34]. A total of 17 *PeMADS* genes were further corrected or completed (Table 2). Among them, 13 *PeMADS* genes lacked the M- or K-domain, while four other genes have incorrect sequences in their open reading frame regions. After correction, the *PeMADS* sequences' quality and integrity were greatly improved. For instance, a K-domain, which is one of the most conserved domains in MADS-box genes, was added to *PeMADS23* (Additional file 2: Figure S1). These 17 full-length *PeMADS* genes were also experimentally cloned from *P. edulis* cDNA with gene-specific primers

**Table 1** List of 42 MADS-box genes identified in *P. edulis* and their homologs in *Oryza* (aa, amino acids)

Name	Gene ID	Homologous	Length (aa)	Intron	Gene location	Subfamily	Predicted gene type
<i>PeMADS1</i>	PH01002127G0260	<i>OsMADS3</i>	266	7	PH01002127:173109–184,410(– stand)	Type II	AG-like
<i>PeMADS2</i>	PH01000306G0610	<i>OsMADS18</i>	257	7	PH01000306:378556–403,756(– stand)	Type II	FUL3 AP1 SQUA-like
<i>PeMADS3</i>	PH01000606G0250	<i>OsMADS14</i>	201	6	PH01000606:177440–184,284(+ stand)	Type II	FUL1 AP1 SQUA-like
<i>PeMADS4</i>	PH01000466G0340	<i>OsMADS68</i>	370	10	PH01000466:212382–216,357(– stand)	Type II	MIKC*
<i>PeMADS5</i>	PH01000437G0930	<i>OsMADS22</i>	228	7	PH01000437:618016–625,547(– stand)	Type II	SVP-like
<i>PeMADS6</i>	PH01000059G1270	<i>OsMADS56</i>	222	6	PH01000059:749594–811,731(+ stand)	Type II	TM3-like
<i>PeMADS7</i>	PH01000759G0450	<i>OsMADS56</i>	226	6	PH01000759:309532–372,975(+ stand)	Type II	TM3-like
<i>PeMADS8</i>	PH01000038G1550	<i>OsMADS47</i>	230	7	PH01000038:1113162–1,119,265(– stand)	Type II	SVP-like
<i>PeMADS9</i>	PH01000317G0080	<i>OsMADS2</i>	209	6	PH01000317:51918–54,493(+ stand)	Type II	PI GLO-like
<i>PeMADS10</i>	PH01001878G0200	<i>OsMADS4</i>	209	6	PH01001878:118831–120,592(+ stand)	Type II	PI GLO-like
<i>PeMADS11</i>	PH01000338G0120	<i>OsMADS31</i>	226	4	PH01000338:109773–112,174(+ stand)	Type II	GGM13-like
<i>PeMADS12</i>	PH01000616G0020	<i>OsMADS33</i>	190	5	PH01000616:15256–18,269(– stand)	Type II	AGL12-like
<i>PeMADS13</i>	PH01002743G0050	<i>OsMADS18</i>	257	7	PH01002743:23635–31,559(– stand)	Type II	FUL3 AP1 SQUA-like
<i>PeMADS14</i>	PH01000006G0210	<i>OsMADS20</i>	281	6	PH01000006:135925–141,334(– stand)	Type II	SQUA-like
<i>PeMADS15</i>	PH01005177G0100	<i>OsMADS26</i>	211	5	PH01005177:60648–64,034(+ stand)	Type II	AGL12-like
<i>PeMADS16</i>	PH01000794G0210	<i>OsMADS4</i>	273	5	PH01000794:136094–138,076(– stand)	Type II	PI GLO-like
<i>PeMADS17</i>	PH01001113G0580	<i>OsMADS32</i>	301	6	PH01001113:351365–358,010(– stand)	Type II	PI GLO-like
<i>PeMADS19</i>	PH01001303G0110	<i>OsMADS37</i>	183	5	PH01001303:60303–74,009(– stand)	Type II	OsMADS37-like
<i>PeMADS20</i>	PH01001952G0230	<i>OsMADS1</i>	244	7	PH01001952:162071–190,513(– stand)	Type II	AGL2-like SEP
<i>PeMADS21</i>	PH01000117G1210	<i>OsMADS37</i>	158	2	PH01000117:863423–865,146(– stand)	Type II	OsMADS37-like
<i>PeMADS23</i>	PH01002755G0230	<i>OsMADS50</i>	215	6	PH01002755:82368–142,165(– stand)	Type II	TM3-like
<i>PeMADS24</i>	PH01000080G0200	<i>OsMADS55</i>	259	7	PH01000080:117443–126,410(– stand)	Type II	SVP-like
<i>PeMADS25</i>	PH01003178G0150	<i>OsMADS65</i>	205	4	PH01003178:58826–64,955(– stand)	Type II	OsMADS37-like
<i>PeMADS26</i>	PH01000222G1060	<i>OsMADS34</i>	216	6	PH01000222:703088–717,231(+ stand)	Type II	AGL2-like SEP
<i>PeMADS28</i>	PH01000222G1190	<i>OsMADS15</i>	192	6	PH01000222:177607–183,539(+ stand)	Type II	FUL2 AP1 SQUA-like
<i>PeMADS29</i>	PH01001278G0330	<i>OsMADS3</i>	229	6	PH01001278:214488–227,311(+ stand)	Type II	AG-like
<i>PeMADS31</i>	PH01021006G0010	<i>OsMADS3</i>	229	5	PH01021006:308–2511(+ stand)	Type II	AG-like
<i>PeMADS32</i>	PH01003236G0170	<i>OsMADS29</i>	251	4	PH01003236:150540–151,005(– stand)	Type II	GGM13-like
<i>PeMADS33</i>	PH01001174G0480	<i>OsMADS5</i>	228	5	PH01001174:303844–317,831(+ stand)	Type II	AGL2-like SEP
<i>PeMADS34</i>	PH01002152G0120	<i>OsMADS56</i>	221	6	PH01002152:73202–96,884(+ stand)	Type II	TM3-like
<i>PeMADS35</i>	PH01000107G0570	<i>OsMADS56</i>	152	4	PH01000107:455631–474,775(– stand)	Type II	TM3-like
<i>PeMADS40</i>	PH01001750G0200	<i>OsMADS58</i>	230	5	PH01001750:217480–232,134(+ stand)	Type II	AG-like
<i>PeMADS41</i>	PH01001188G0490	<i>OsMADS15</i>	259	7	PH01001188:311503–334,928(+ stand)	Type II	FUL2 AP1 SQUA-like
<i>PeMADS43</i>	PH01000077G1380	<i>OsMADS55</i>	224	6	PH01000077:896999–903,795(+ stand)	Type II	SVP-like
<i>PeAP<sub>3</sub></i>	N/A	<i>OsMADS16</i>	200	7	PH01001272:176122–180,192(+ stand)	Type II	AP <sub>3</sub> GLO-like
<i>PeAGL6</i>	N/A	<i>OsMADS6</i>	240	4	PH01002217: 12110–41,791(+ stand)	Type II	AGL6-like
<i>PeMa1</i>	N/A	<i>OsMADS73</i>	97	0	PH01010177:5039–5332(+ stand)	Type I	Ma
<i>PeMa2</i>	N/A	<i>OsMADS73</i>	98	0	PH01000557:531750–532,199(– stand)	Type I	Ma
<i>PeMa3</i>	N/A	<i>OsMADS78</i>	212	0	PH01000604:455236–455,874(– stand)	Type I	Ma
<i>PeMa4</i>	PH01000954G0350	<i>OsMADS72</i>	377	1	PH01000954:196706–198,735(– stand)	Type I	Ma
<i>PeMa5</i>	N/A	<i>OsMADS75</i>	184	0	PH01002999: 94340–94,894(+ stand)	Type I	Ma
<i>PeMa6</i>	N/A	<i>OsMADS73</i>	87	0	PH01224566:45–309 (– stand)	Type I	Ma



to further confirm the sequence accuracy (Additional file 1: Table S1). To determine the accuracy levels of the modified sequences, we chose one gene, *PeMADS23*, to undergo a functional analysis. Using a subcellular localization prediction tool, *PeMADS23* (corrected sequence) was found to have a nuclear subcellular location signal that is not found in the original sequence PH01002755G0230. In the experimental validation, the *PeMADS23*-YFP fusion protein was observed to be localized in the nucleus of the *Arabidopsis* protoplast,

**Table 2** The detail information about these 17 genes before and after bioinformatics corrections

Gene ID	Incomplete Reason	Original Length (aa)	Corrected Length (aa)
<i>PeMADS13</i>	Lack of K-domain	124	257
<i>PeMADS19</i>	Lack of K-domain	132	183
<i>PeMADS20</i>	Lack of K-domain	79	244
<i>PeMADS21</i>	Lack of K-domain	152	258
<i>PeMADS23</i>	Lack of K-domain	105	215
<i>PeMADS25</i>	Lack of K-domain	86	205
<i>PeMADS26</i>	Prior 319 bp is APO-domain	160	216
<i>PeMADS28</i>	Lack of M-domain	235	192
<i>PeMADS29</i>	Lack of M-domain	200	229
<i>PeMADS31</i>	Lack of M-domain	152	229
<i>PeMADS32</i>	M-domain is truncated	179	251
<i>PeMADS33</i>	Prior 184 bp is incorrect	181	228
<i>PeMADS34</i>	Prior 100 bp is incorrect	202	221
<i>PeMADS35</i>	Lack of M-domain	309	152
<i>PeMADS40</i>	Lack of M-domain	273	230
<i>PeMADS41</i>	Lack of M-domain	173	259
<i>PeMADS43</i>	Lack of M-domain	165	224

while the PH01002755G0230-YFP fusion protein was observed in the cytoplasm (Fig. 2). Thus, the experimental results corroborated the predictions. Additionally, the K-domain appears to be essential for a MADS-box gene to gain entry into the nucleus. Compared with those of the BambooGDS, the sequences obtained using our protocols are more accurate, which will assist further MADS-box gene functional analyses.

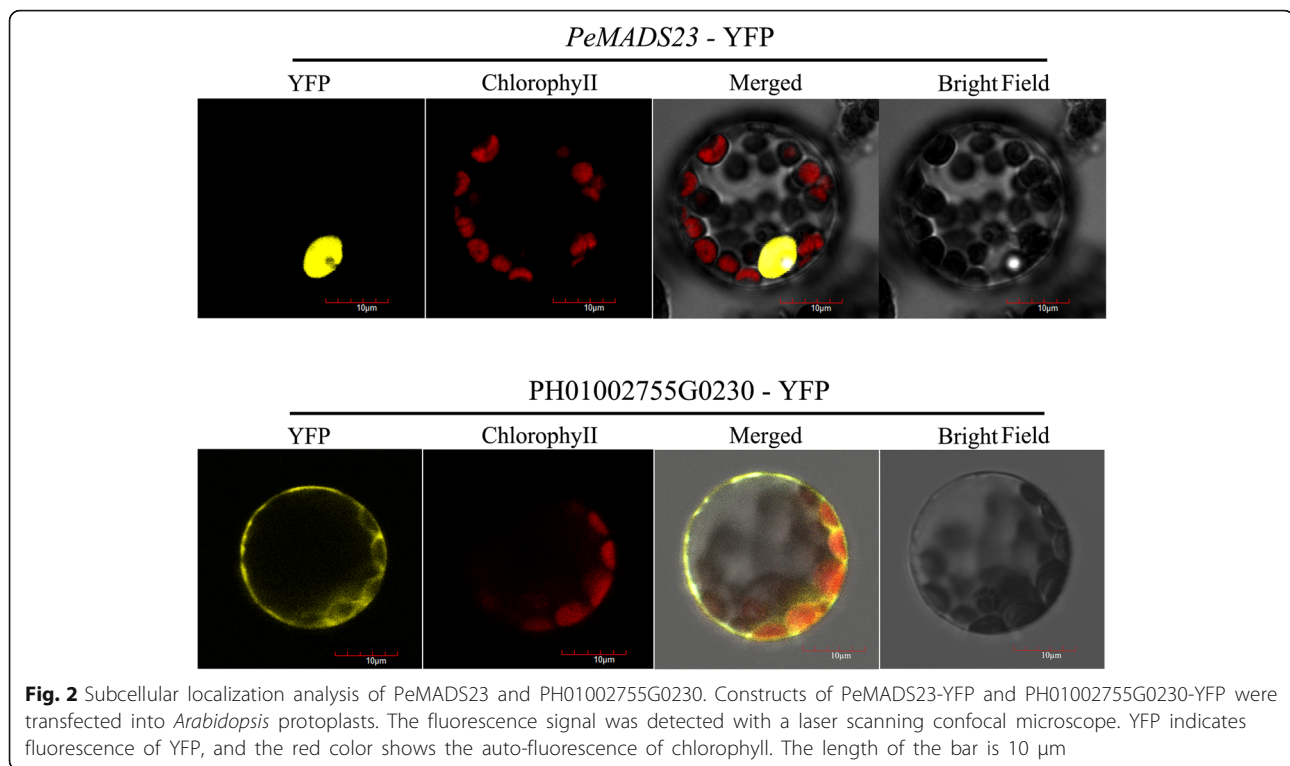
#### Phylogenetic and conserved motif analyses of *PeMADS*s

A phylogenetic tree of both type-I and type-II MADS-box genes was constructed to determine the evolutionary relationships between *P. edulis* MADS-box genes and the known MADS-box genes of *Arabidopsis* and *Oryza*. The exhibited groupings of MADS genes in moso bamboo were similar to those of the established model plants and can be further divided into 16 clades (Fig. 3). *PeMADS*s, except for the *AGL17*, *FLC*-like, and *Mβ*- and *My* subgroups, formed 11 subgroups. Most *PeMADS*s belong to the MIKC<sup>C</sup> type: two *P. edulis* orthologs each were present for the *Arabidopsis* GGM13 and *AGL12* clades; three were present for the *AGL2*(*SEP*) and *OsMADS37* clades; four were present for the *AG* and *Solanum tuberosum* MADS11(*SVP*) clades; five were present for *GLO* (PI)-like, and six were present for *Ma*, *TM3*, and *SQUA*-like clades. Only one MIKC\*-type gene, *PeMADS4*, has been identified in *P. edulis*. It has a longer I-domain, which appears to be the most prominent characteristic of the MIKC\*-type proteins [49].

The motif distribution in *PeMADS* proteins was analyzed using the MEME program. The MEME software identified 20 conserved motifs in *PeMADS*, as well as their distribution (Fig. 4). The detected motifs were annotated using SMART protein analyzing software. Motifs 1 and 5 contain the typical MADS-box domain. Motifs 2 and 4 contain the second most conserved signature motif, the K-domain. Motifs 3, 6 and 18 contain the I-domain. C-terminals were the least conserved region of the MIKC proteins and had the most unknown motifs.

#### Long-terminal repeat-retrotransposon (LTR-RT) analysis in long introns of *PeMADS*s

The CDSs and corresponding genomic DNA sequences of the *PeMADS*s were used to identify the gene structures of the *PeMADS* genes, with four (*PeMADS1*, -2, -6, and -7) having extraordinary long introns. In fact, those of *PeMADS6* and -7 are ~60 kb. We analyzed potential repeats in the *PeMADS* gene family using Repeat Masker online software [50] (Additional file 1: Table S3). The four above-mentioned MADS-box genes have a number of repeats in their intron areas. After analyzing these repetitive sequences, 140 transposable elements (TEs) were identified (Additional file 1: Table



**Fig. 2** Subcellular localization analysis of PeMADS23 and PH01002755G0230. Constructs of PeMADS23-YFP and PH01002755G0230-YFP were transfected into *Arabidopsis* protoplasts. The fluorescence signal was detected with a laser scanning confocal microscope. YFP indicates fluorescence of YFP, and the red color shows the auto-fluorescence of chlorophyll. The length of the bar is 10 µm

S4). LTR-RTs were the most abundant class, with 44 LTR/*Gypsy* and 26 LTR/*Copia* (26) superfamily members. In addition to LTR-RTs, DNA transposons, including DNA/*hAT-Tag1*, DNA/*Merlin*, and DNA/*Harbinger*, were found in these MADS-box genes.

#### **cis-element analysis of MADS-box family gene promoters and intron sequences**

Most of the well-characterized MADS-box genes contribute to the processes of plant growth and responses to hormones or environmental stimuli, such as photoperiod and temperature. Determining the promoter region features of *PeMADS*s will help us to understand the expression patterns of bamboo MADS-box genes. To identify *cis*-regulatory elements in *PeMADS* genes, we extracted the promoter and intron regions for each *PeMADS* gene from the *P. edulis* genome and analyzed them using the PlantCare server. We categorized all of the *cis*-elements into nine broad categories based on their responsiveness to environmental stimuli. Details of the *cis*-elements are found in Additional file 1: Table S6. In the pie chart (Fig. 5), the proportion of light-responsive elements is the greatest, followed by tissue-specific, plant hormone-responsive, abiotic stress-responsive, and core promoter elements. G-box, Sp1, and TCT-motifs were the most frequently identified light-responsive elements. We identified several hormone-responsive *cis*-elements, such as ABRE, CGTCA-motif, TGACG, and TCA-elements, and abiotic stress-responsive elements, such as ARE,

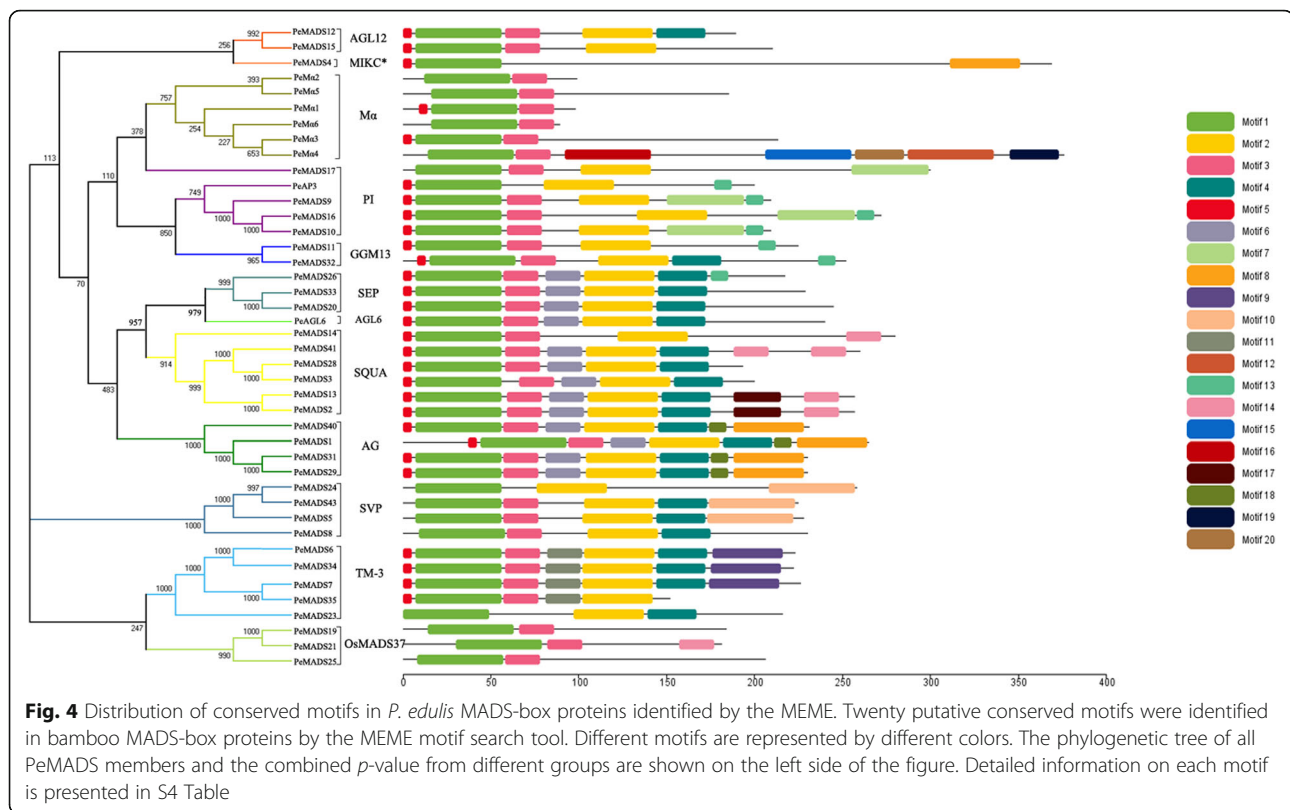
AAGAA-motif, LTR, HSE, GC-motif and TC-rich repeats. Most genes contain the endosperm expression elements Skn-1 and GCN4-motif. Some contain the meristem tissue-specific elements CAT and CCGTCC-box. Furthermore, we found some protein-binding sites in *PeMADS* promoter regions. The most numerous binding sites are MBS, CarG, and CCAAT-box. Previous studies showed that MADS-box proteins regulate the expression of target genes by binding to CarG motifs in their promoter regions [51]. CarG-motifs are not specific to promoter regions, but can also occur in the intron area. For instance, the *Arabidopsis* *AG* gene has a CarG-motif in its 1st intron [52, 53]. In our study, CarG motifs existed many times throughout *PeMADS* promoters and introns (Additional file 1: Table S7). The presence of these core binding sites indicated that *PeMADS* proteins form either homodimers or heterodimers to bind to the CarG-box sequence and affect the floral transition. The presence of the plant AP2-like-binding element implies that the three genes *PeMADS13*, *24*, and *43* can respond to endogenous floral developmental signals. This prediction is consistent with *PeMADS13*, *24*, and *43* showing relatively greater expression levels in floral tissues than in leaf tissues.

#### **Expression analysis of MADS-box genes during moso bamboo floral development**

We investigated the spatial expression patterns of the MADS-box genes in bamboo leaf tissues and in the different stages of floral development. *PeMADS*s were



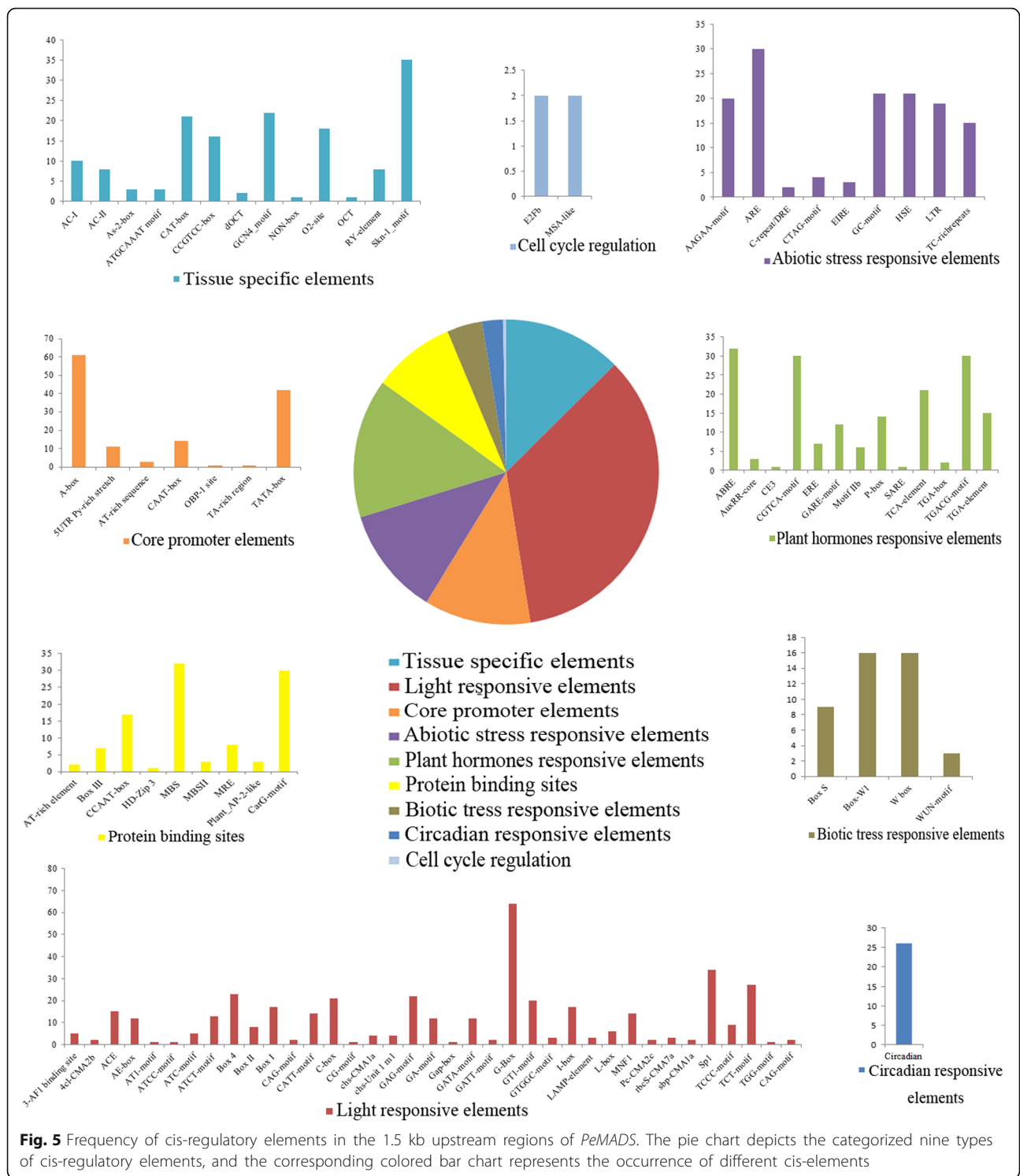




Bamboo inflorescence development can be divided into four stages, F1–F4, from first floral bud to anthesis. *PeMADS14*, *-15*, *-19*, *-21*, *-24*, *-28*, and *-40* were mainly expressed at the F1 stage. *PeMADS2*, *-5*, *-9*, *-10*, *-16*, *-20*, and *-29* were prominently expressed at the F2 and F3 stages. Then, the expression levels of *PeMADS1*, *-3*, *-9*, *-25*, *-32*, and *-33* were greater higher in the F4 stage. In general, bamboo MADS-box genes within subfamilies share relatively similar expression patterns. Based on the expression patterns (Fig. 6), MADS-box genes, such as *SQUA*-like (*PeMADS13*, *-14*, and *-28*) and *OsMADS37*-like genes (*PeMADS19* and *-21*) were found to be involved in floral meristem activity and were highly expressed at the F1 stage. The “ABCDE” model genes *AGL2*-like (*PeMADS20* and *-33*), *GLO*-like (*PeMADS9*, *-10*, and *-16*), and *AG*-like (*PeMADS1*, *-29*, and *-31*) were up-regulated during floral-organ development. Homologs of the *SVP/StMADS11*-like subfamily are hallmark meristematic genes, and *PeMADS5*, *-8*, *-24*, and *-43* displayed their greatest expression levels in F2, leaf, F1, and F4, respectively. The type-I genes *PeMa1*, *-3*, *-4*, and *-5* were highly expressed during the late developmental stages of the bamboo inflorescence. *PeMADS5* was greatly expressed in the F2 stage, and its expression decreased with the development of the bamboo inflorescence, suggesting a role for this gene in spikelet initiation.

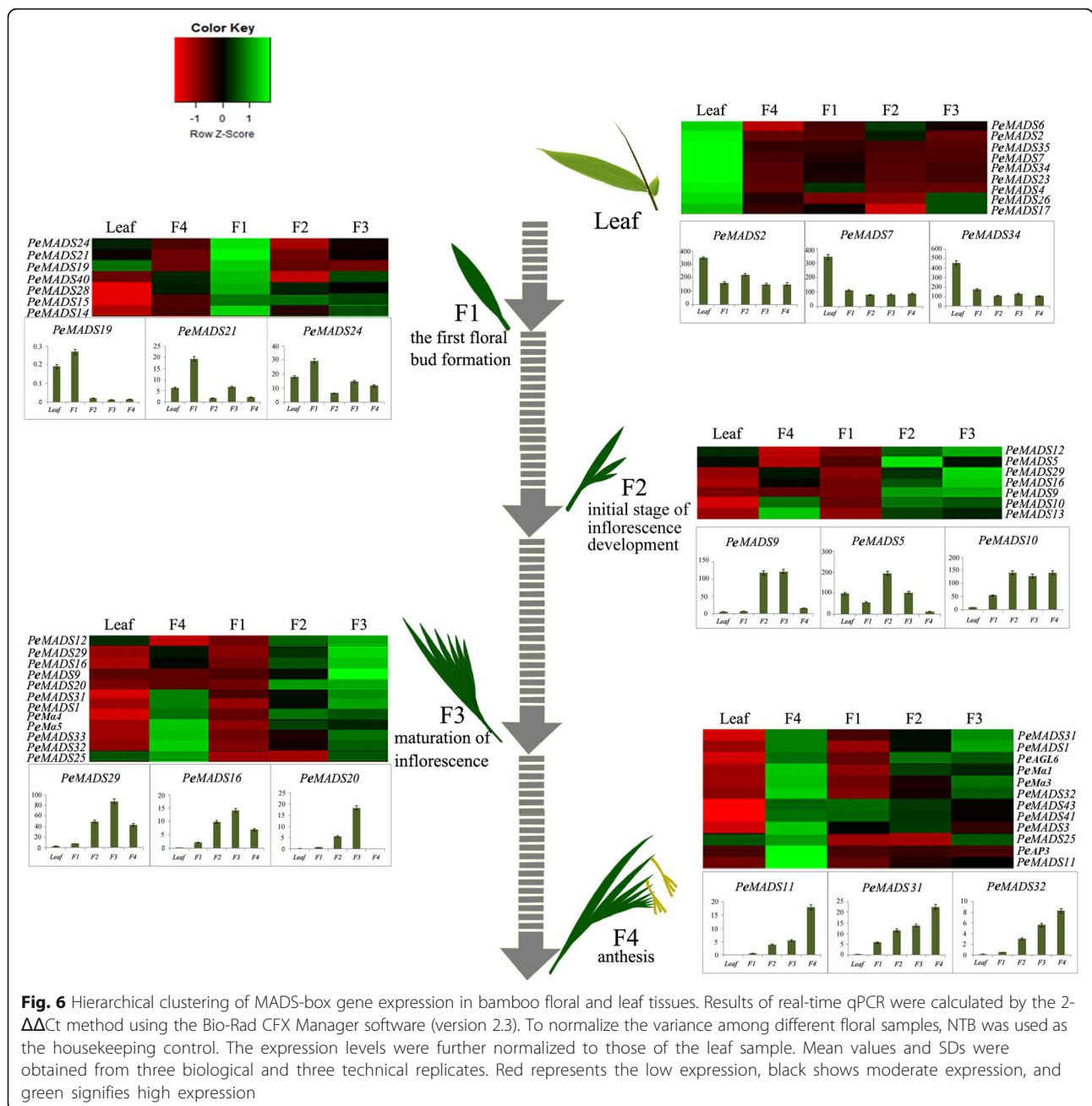
#### Overexpression of *PeMADS5* in wild-type *Arabidopsis* plants causes an early flowering time and abnormal floral morphology

Genes from *SVP/StMADS11*-like subgroup are involved in panicle branching, flowering time determination, and floral meristem specification in *Arabidopsis* and *Oryza*. In this study, we found four *SVP/StMADS11*-like genes, *PeMADS5*, *-8*, *-24*, and *-43*, in bamboo. Among them, *PeMADS5* was expressed highly during inflorescence initiation. We selected this gene for the functional studies. We obtained 12 single-copy independent transgenic lines for *PeMADS5*. Flowering time was determined in the T3 generation under the long-day conditions, and *35S::PeMADS5* plants flowered significantly earlier than wild-type plants (Fig. 7a and b). Four of the transformed *Arabidopsis* T3 lines (*35S::PeMADS5* #2, #4, #5, and #10) showed various degrees of phenotypic alterations in their reproductive organs compared with wild-type plants (Fig. 8 and Additional file 4: Figure S3). To confirm the overexpression of transgenes, *PeMADS5* transcript levels were analyzed by RT-PCR in these four transgenic lines (Fig. 7c). Relative *PeMADS5* expression levels in *35S::PeMADS5* #2 and #10 were greater than in other lines, and they also had ‘strong’ phenotypes as shown in Fig. 8e, f. Flowers of *35S::PeMADS5* #2 and #10 exhibited sepals that formed leaf-like structures and did not completely enclose the inner developing organs



(Fig. 8f and j). Additionally, 35S: *PeMADS5* #2 had five petals instead of four (Fig. 8g). Furthermore, sepals remained attached in fruit of both transgenic lines #2 and #10 (Fig. 8h and l), which was in contrast to wild-type *Arabidopsis* plants (Fig. 8d). In addition to the phenotypic analysis, the expression levels of some

flowering-related genes, including those involved in flowering time and flower organ development, were also analyzed (Fig. 8m and n). In both rosette leaves and inflorescences, *AGL24* and *SOC1* transcripts were significantly upregulated in 35S::*PeMADS5* #2 and #10 compared with wild-type, while *API* was downregulated.



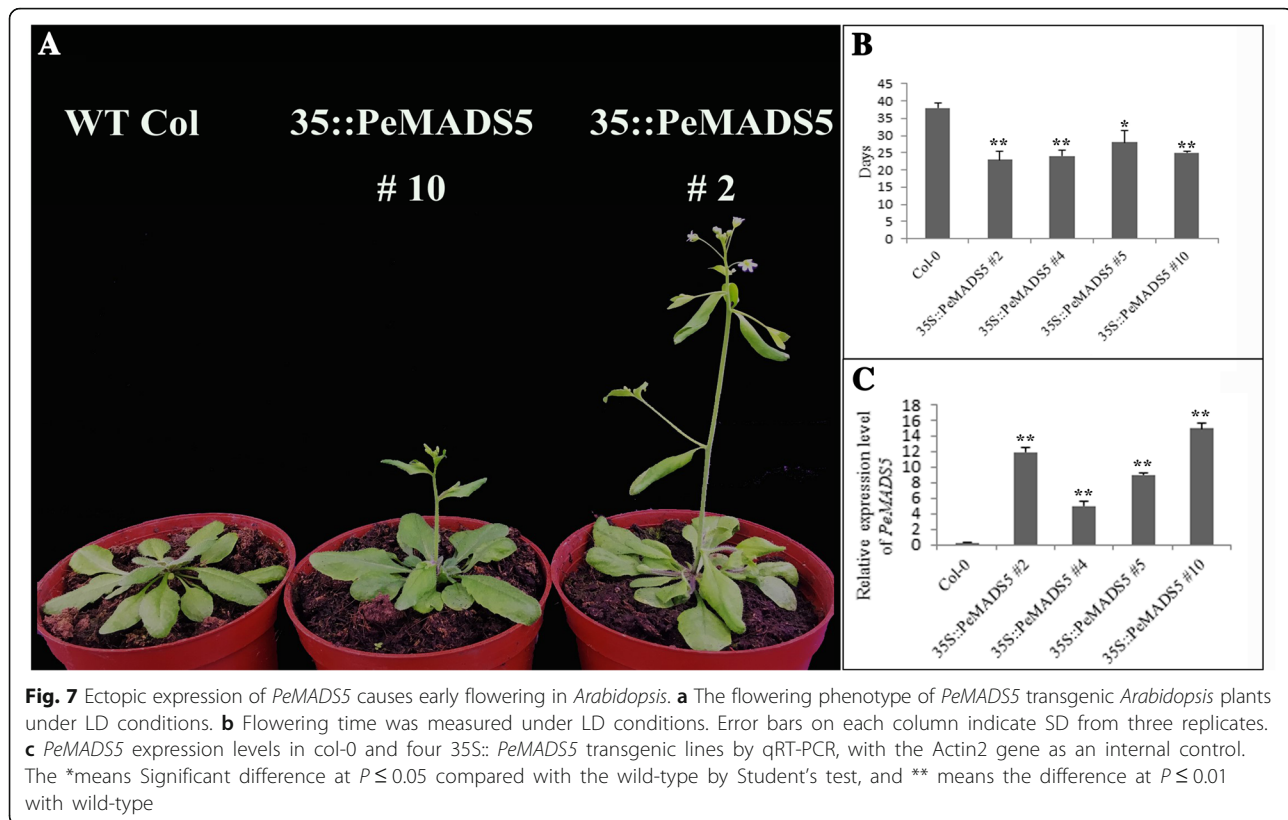
**Fig. 6** Hierarchical clustering of MADS-box gene expression in bamboo floral and leaf tissues. Results of real-time qPCR were calculated by the 2- $\Delta\Delta C_t$  method using the Bio-Rad CFX Manager software (version 2.3). To normalize the variance among different floral samples, NTB was used as the housekeeping control. The expression levels were further normalized to those of the leaf sample. Mean values and SDs were obtained from three biological and three technical replicates. Red represents the low expression, black shows moderate expression, and green signifies high expression

The expression levels of *FT* and *SEP3* in 35S::*PeMADS5* were nearly 0.3- and 0.5-fold, respectively, those of the wild-type inflorescences, while they were expressed at slightly different levels in leaves. Thus, *PeMADS5* may act as a floral activator in bamboo flowering.

#### Interaction of *PeMADS5* with other floral-related MADS-box proteins

In *Arabidopsis*, AGL24 interacts with SOC1 and FUL in the shoot apical meristem to promote flowering [54]. During floral organ formation, AGL24–AP1 dimers can also

interact with the class B gene *PISTILLATA* (*PI*) and the class E gene *SEPALLATA* (*SEP*) [55, 56]. Therefore, we performed a yeast two-hybrid assay using the protein-coding regions of *PeSOC1* (*PeMADS34*), *PeAP1* (*PeMADS2*), *PeSEP3* (*PeMADS20*), and *PePI* (*PeMADS16*) to obtain insights into the interaction patterns of the *PeAGL24* (*PeMADS5*) protein in bamboo. The coding sequences of *PeMADS2*, -5, -16, -20, and -34 were fused to the binding and activating domains and their abilities to interact were determined (Fig. 9). *PeMADS5* (AGL24) could interact independently with *PeMADS34* (SOC1) and



PeMADS2 (AP1), but not with PeMADS16 (PI) or PeMADS20 (SEP3).

## Discussion

### A practical alternative way to obtain a complete gene sequence

Owing to the current assembly of the bamboo genome, some *PeMADS* sequences were incorrectly annotated and lack full-length CDSs. For instance, some MADS-box genes identified in the BambooGDS lack the characteristic M- or K-domain. In this study, we employed a transcriptome assembling and polishing method to achieve more accurate MADS-box sequences in bamboo. Briefly, we first used the truncated *PeMADS* sequences as query seeds to seek optimally aligned short reads. These reads extended both the 5'- and 3'-end regions until full-length CDSs with 5' UTRs and 3' UTRs were achieved. By cloning full-length genes and validating protein subcellular localizations, we demonstrated that this bioinformatics procedure reliably produced full-length MADS-box gene sequences.

### MADS-box type-I subfamily members may be lost during *Phyllostachys*' evolution

A total of 36 MIKC-type MADS-box genes in moso bamboo were identified in this study. This number is similar to that in *Arabidopsis* (39), *Oryza* (38), and other

well-studied species (Table 3). Non-candidate genes belonging to the type-I MADS-box were found. Using *Oryza* and *Brachypodium* type-I genes sequences as query, only six Ma $\alpha$  genes have been found. In gymnosperms, type-I MADS-box genes are underrepresented [57]. Only two type-I MADS-box genes are present in the most recent common ancestor of seed plants [58]. Although the number of type-I MADS-box genes appeared to have dwindled in some angiosperms, including wheat, poplar, and barley, we observed an extreme decrease in bamboo [46, 59, 60]. *Brachypodium* has the same number of type-II genes as bamboo but also has fewer members of the type-I subfamily (Table 3). Therefore, the common ancestor of bamboo and *Brachypodium* may have undergone a gene loss event after it initially diverged from *Oryza* 45–50 million years ago. Subsequently, we hypothesize that bamboo experienced a second gene loss event after speciation. The expansion and construction of orthologous clusters showed the greatest number, among the grass family members, in bamboo after species divergence [61]. Thus, gene families in bamboo are more divergent and polarized than in other grass species. Another hypothesis for the lack of type-I genes is based on numerous scaffolds not being obtained by the *Phyllostachys* whole-genome shotgun sequencing, resulting in the lack of assembled pseudochromosomes. Given the presence of small fragments in the

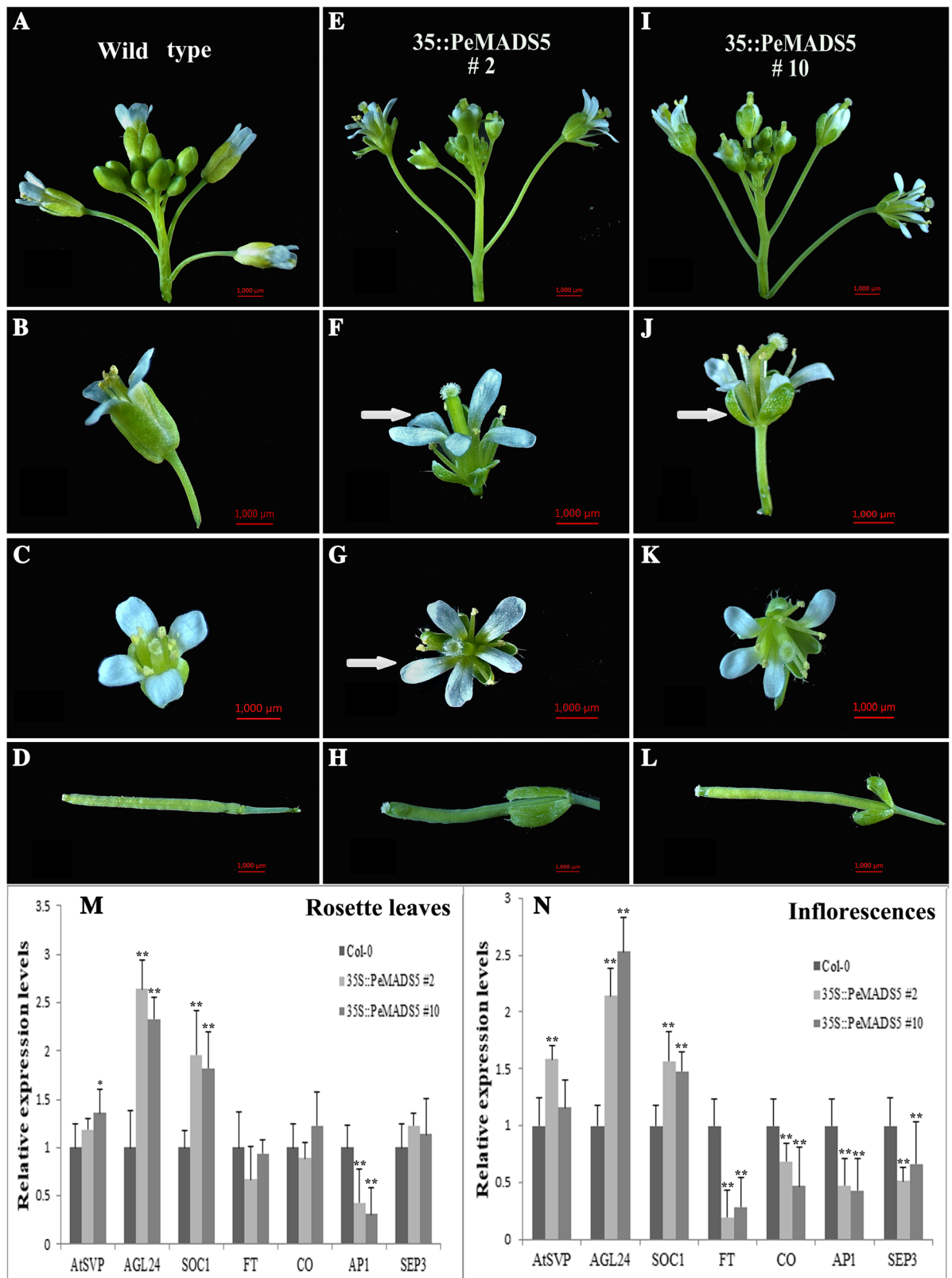


Fig. 8 (See legend on next page.)

(See figure on previous page.)

**Fig. 8** Morphological analysis of 35S::PeMADS5 transgenic *Arabidopsis*. Light micrographs of Wild-type *Arabidopsis* Col-0 (a–d) and 35S::PeMADS5 transgenic plants (#2: E to H; #10: I to L). **a** Wild-type *Arabidopsis* Col-0 inflorescence. **b** Side view and **(c)** top view of a wild-type *Arabidopsis* Col-0 flower at anthesis. **d** A wild-type *Arabidopsis* matured fruit. **e** 35S::PeMADS5 #2 **(f)** 35S::PeMADS5 #10 inflorescences. **f** Side view and **(g)** top view of a 35S::PeMADS5 #2 flower at anthesis. Note the extra petal indicated with an arrow. **j** Side view and **(k)** top view of a 35S::PeMADS5 #10 flower at anthesis. Note the sepals with leaf-like characteristics and did not completely enclose the inner developing organs. Both 35S::PeMADS5 #2 **(h)** and #10 **(l)** matured fruits have attached petals denoted with arrows. Relative expression levels of different flowering time genes (*SVP*, *AGL24*, *SOC1*, *FT*, *CO*, *API* and *SEP3*) in both leaves **(m)** and flowers **(n)** as determined by qRT-PCR, with the Actin 2 gene as an internal control. Error bars on each column indicated from three replicates. The \* indicates Significant difference at  $P \leq 0.05$  compared with the wild-type by Student's test, and \*\* means the difference at  $P \leq 0.01$  with wild-type

assembly, the lack of type-I subfamily members may be a statistical riddle because of the imperfect genome sequence [61]. Conversely, orthologous type-I genes may exist in bamboo but were not identified because of the limited number of phylogenetically informative sites. In summary, bamboo may have a unique type-I subfamily evolutionary pattern but share a common type-II subfamily evolutionary pathway with other angiosperms.

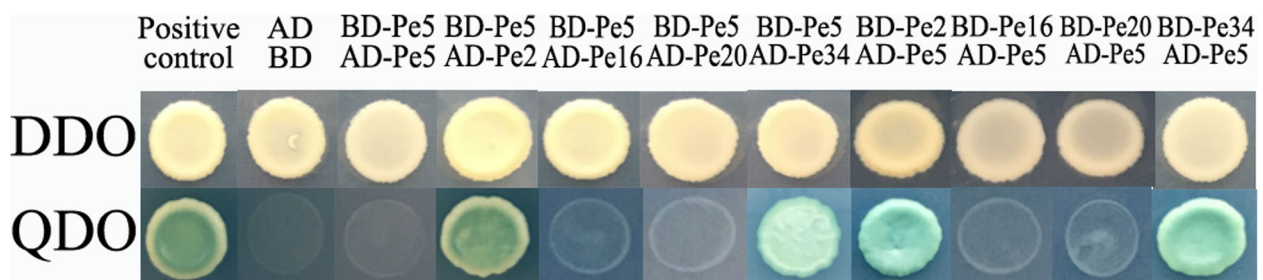
#### Presence of long introns and the insertion of LTR-RT elements in the *Phyllostachys* MADS-box family

A notable characteristic of the structure of bamboo MADS-box genes is the presence of long introns. This is similar to findings in the repeat-rich genomes of *Vitis vinifera*, *Zea mays*, and Norway spruce [62–64]. An analysis of sequences containing long introns revealed that they all contain numerous repeats, suggesting that intron expansion in bamboo was promoted by repeat insertion. By contrast, exon size and the number of MADS-box genes were consistent with those of other well-studied species. Long introns did not influence expression levels (Fig. 6). TEs are the most abundant DNA components in higher eukaryotes [65], and is approximately 59% of the bamboo genome consists of TEs [61]. In the full-length genome sequences of *PeMADS1*, –2, –6, and –7, 140 TEs were found. Most of the repeats could be assigned to known TE repeat families (Additional file 1: Table S5). LTR-RTs were the most abundant fraction of TEs, with the *Gypsy*-type LTR superfamily members

being more abundant than *Copia*-type LTR superfamily members. This was consistent with the genome-wide analysis of LTR-retroelements in *P. edulis* [66]. These TEs are prone to insertion and/or maintenance in the intronic regions of MADS-box genes. Thus, *PeMADS*s, with their abundant repetitive sequences, may be hot-spots for TE insertion. Large-genome species, such as maize, barley, and wheat, contain a large number of LTR-RT element that, which have been amplified within the last few million years [67, 68]. Therefore, we hypothesized that the accumulation of LTR-RTs in bamboo gene families may be a main factor that led to the increased genome size without polyploidization, like that identified in *Oryza australiensis*.

#### Expression analysis of *Phyllostachys* MADS-box genes and their orthologs during inflorescence development

To gain insight into bamboo MADS-box gene expression patterns, we carried out a qPCR expression study using different floral tissues. The study revealed that critical regulatory genes for spikelet meristem development might be conserved between *Phyllostachys* and *Oryza*. For example, genes from the *AG*-like subfamily are essential for reproductive structure morphogenesis [69–71]. As shown in Fig. 6, bamboo *AG*-like genes *PeMADS1*, –29, and –31 are primarily expressed in the later stages of floral development, whereas their *Oryza* orthologs *OsMADS3* and –58 have relatively greater expression levels in reproductive tissues [18]. Several



**Fig. 9** Yeast two-hybrid assays of PeMADS5 and its interacting proteins. The *PeMADS5*, *PeMADS2*, *PeMADS16*, *PeMADS20* and *PeMADS34* genes were fused both in-frame to the GAL4 DNA-binding domain (BD)-coding sequences and activation domain (AD)-coding sequences. Cell growth on -Leu-Trp dropout selective medium (-DDO) represents normal cells (upper panel) and -Leu-Trp-His-Ade dropout selective medium (-QDO) represents positive interactions (bottom panel). Pe refers to PeMADS

**Table 3** Total number of MADS-box genes within each group among nine species

	Ma	Mβ	My	I type	MIKC <sup>C</sup>	MIKC <sup>*</sup>	II type	Pseudogenes	Total
<i>P. edulis</i>	6			6	36	1	36	8	42
<i>Arabidopsis</i>	25	20	16	61	39	8	45		107
<i>Oryza</i>	13	9	10	32	40	3	43		75
Maize	27	3	2	32	39	4	43		75
<i>Sorghum</i>	26	2	2	30	33	2	35		65
<i>Physcomitrella patens</i>	2	2	3 (unclassified)	7	6	11	17	2 (1 Mβ; 1 MIKC <sup>*</sup> )	26
<i>Brachypodium</i>	9	7	2	18	32	7	39		57
Poplar	23	12	6	41	55	2		7 (Mδ)	105
Soybean	37	14	24	75	81	7	88		163

*SQUA*-like genes are proposed to be vital regulators in inflorescences or floral meristems and are involved in organ identity [72]. In the present study, *Phyllostachys SQUA*-like genes *PeMADS2*, *-3*, *-13*, and *-41* were significantly expressed at all stages of inflorescence development, with expression patterns parallel to those of their putative *Oryza* orthologs *OsMADS14*, *-15*, and *-18*, respectively [18]. Nevertheless, modifications in orthologous gene expression patterns may be necessary owing to neofunctionalization and subfunctionalization. *OsMADS34*, a member of the *SEP* subgroup, was a key regulator of rice inflorescence and spikelet architecture. In rice, *OsMADS34* expression was detected in the floral meristem, and the *osmads34* mutants develop altered inflorescence morphology [73]. Additionally, *PeMADS26*, the *OsMADS34* homolog in bamboo, was highly expressed in leaves but not florets. Similarly, *PeMADS6* and *-34* from the TM3-like subgroup were predominantly expressed in the bamboo leaf but not in the spikelet meristems. In *Oryza*, *MADS56* showed a high transcript accumulation level at all stages, especially in panicle materials [18]. We surmised that the differences in inflorescence structures might be the result of expression pattern changes in the conserved genes of the grass family.

#### ***PeMADS5* may interact with *PeMADS34* and *PeMADS2* to control flowering time and floral organ development**

*PeMADS5* belongs to the *StMADS11/SVP* clade of the MADS-box gene family, which has two diversified functional members in *Arabidopsis*: *SVP* and *AGL24* [74]. In *Arabidopsis*, *AGL24* acts as an integrator of multiple flowering signals and function together with *SOC1* to regulate the floral transition and inflorescence meristem identity [75, 76] *SVP* controls flowering time by negatively regulating the expression of *FLOWERING LOCUS T (FT)* by directly binding to the CArG motifs in the *FT* sequence [77]. The overexpression of *AGL24* in *Arabidopsis* results in early flowering and floral abnormalities, such as leaf-like sepals, or the transformation of floral meristems into inflorescence meristems [75, 76, 78]. In contrast, the overexpression of *SVP* results in late flowering and the loss of

carpels, as well as the conversion of flowers into shoot-like structures [78]. The ectopic overexpression of *PeMADS5* in *Arabidopsis* triggered an earlier flowering time in all of the transgenic lines, and 4 out of the 15 lines developed aberrant flower phenotypes (Fig. 8). The expression level of *PeMADS5* was greater in transgenic lines with an obvious mutant phenotype, and proportionately lower in lines with less obvious mutant phenotypes (Fig. 7c). The phylogenetic analysis of *PeMADS5* revealed that it had a closer relationship to *OsMADS22* and *OsMADS55* (Fig. 3). The constitutive expression of *OsMADS22* and *OsMADS55* leads to floral reversion phenotypes, including leaf-like sepals, which is similar to the *PeMADS5*-overexpression phenotype, but *OsMADS55* acts as a floral repressor by suppressing the expression levels of *Hd3a* and *SOC1* [79]. In *Arabidopsis*, the identity of each floral organ has been determined by a specific combination of floral homeotic genes constituting the ABCDE model, in which the class E *SEP3* and class A *API* genes have critical roles [80–82]. Thus, the relatively low expression levels of *SEP3* and *API* (Fig. 8m and n) may affect normal floral development, leading to the abnormal floral phenotypes of the *35S::PeMADS5* plants. The ectopic expression of class C (*AG* subfamily) in *Arabidopsis* results in the conversion of petals to stamens and of sepals to carpels, as observed in *35S::AG* lines, and this is similar to the *35S::PeMADS5* phenotypes [83]. AP1–SVP and AP1–AGL24 dimers can bind CArG boxes in the second *AG* intron and affect the formation of floral organs [53]. It is likely that multiple mechanisms act on *PeMADS5* through class C, D and E genes. Furthermore, the yeast two-hybrid assay showed that the *PeMADS5* protein interacts with *PeMADS2* (*PeAP1*) and *PeMADS34* (*PeSOC1*) proteins. This suggests that the interaction domains of the *AGL24*, *SOC1*, and *AP1* proteins are conserved between *Arabidopsis* and bamboo. In conclusion, we found that *PeMADS5* is functionally similar to *AGL24*, that it may interact with *PeSOC1* as an integrator of flowering inducers, and that it associates with *PeAP1* to regulate flower development.

## Conclusion

The recent release of the *P. edulis* genome sequence enabled us to identify and comprehensively analyze its MADS-box family. We identified 42 full-length bamboo MADS-box genes, 6 that were type I and 36 that were type II. Consequently, we hypothesized a low formation rate and high destruction rate for type-I genes in *Phyllostachys*. The MADS-box genes and proteins of moso bamboo were classified and their evolutionary relationships with those of other eudicots were evaluated using phylogenetic and structural analyses. A survey of the expression levels of bamboo MADS-box genes in floral and leaf tissues revealed that some MADS-box genes are involved in inflorescence development. Most *Phyllostachys* MADS-box genes appeared to have similar expression patterns to those of their orthologous genes. MADS-box genes with specific expression patterns may play particular functions in bamboo floral development and can be considered candidate genes for cloning and further functional analyses. In addition, the overexpression of *PeMADS5*, a candidate gene from the *StMADS11* clade, in *Arabidopsis* caused early flowering and abnormal floral organ development. Therefore, further studies are required to understand how *PeMADS5* is regulated and whether and how it regulates other genes to control flower development in *Phyllostachys*.

## Additional files

**Additional file 1: Table S1.** List of primers used for *PeMADS* genes cloning. **Table S2.** List of primers used for qPCR in *P. edulis*. **Table S3.** List of repeat numbers in the *PeMADS* gene family identified by the Repeat Masker online software. **Table S4.** List of transposable elements in *PeMADS1*, -2, -6, and -7. **Table S5.** The primers for Real-time quantitatively RT-PCR in *Arabidopsis*. **Table S6.** List of cis-elements in 1.5 kb upstream promoter regions of *PeMADSs*. **Table S7.** List of CArG box in *PeMADSs* promoter and intron region. (XLSX 136 kb)

**Additional file 2: Figure S1.** The alignment of the cloned sequence, *PeMADS23*, and PH01002755G0230. (TIF 1624 kb)

**Additional file 3: Figure S2.** qPCR expression analysis of bamboo MADS-box genes in floral and leaf tissue (TIF 5256 kb)

**Additional file 4: Figure S3.** The phenotypes of *PeMADS5* transgenic *Arabidopsis* plants under LD conditions (A and B) and SD conditions (C). (TIF 2203 kb)

## Abbreviations

AG: AGAMOUS; AGL: AGAMOUS LIKE; AP: APETALA; cDNA: Complementary DNA; CFO1: Chimeric Floral Organs; DEF: DEFICIENS; FLC: Flowering Locus C; FUL: FRUITFULL; GLO: GLOBOSA; HMM: Hidden Markov Model; LTR-RTs: Long Terminal Repeat-Retrotransposons; ME2: Myocyte Enhancer Factor 2; NCBI: National Center for Biotechnology Information; NJ: Neighbor-joining; NTB: Nucleotide Tract-Binding Protein; PCR: Polymerase chain reaction; qPCR: Quantitative Real-time Polymerase Chain Reaction; SEP: SEPALLATA; SHP: SHATTERPROOF; SOC: Suppressor of Overexpression of CONSTANS; SQUA: SQUAMOSA; SRF: Serum Response Factor; SVP: Short Vegetative Phase; TEs: Transposable Elements; TF: Transcription factor

## Acknowledgements

We thank Lesley Benyon, PhD, from Liwen Bianji, Edanz Group China ([www.liwenbianji.cn/ac](http://www.liwenbianji.cn/ac)), for the modification of article language.

## Funding

This study was supported by the National Natural Science Foundation of China (No. 31270715, 31000295), the Ministry of Science and Technology of the People's Republic of China (No. 2012CB723008), Zhejiang Provincial Specialized Research Fund for Bamboo Breeding (No. 2012C12908-2), Top Key Discipline of Forestry of Zhejiang Province [KF201304] and Innovation Program for Graduate Student of Top key Discipline of Forestry.

## Availability of data and materials

All relevant supplementary data is provided within this manuscript as Supplementary files. Genomic sequences and gene annotation information of moso bamboo presented in this report are available in the bamboo genome database (BambooGDB, <http://202.127.18.221/bamboo/index.php>). The *P. edulis* transcriptome reads were downloaded from the NCBI SRA database (Accession: SRR4450542, SRR4450543, SRR4450544, SRR4450545, SRR4450546, SRR4450547, SRR4450548, SRR4450549, SRR4450550, SRR4450551). The full-length sequences of the MADS-box gene family of *A. thaliana* and *O. sativa* are available in the *Arabidopsis* Information Resource (TAIR) (<https://www.arabidopsis.org/index.jsp>) and Rice Genome Annotation Project ([http://rice.plantbiology.msu.edu/home\\_contacts.shtml](http://rice.plantbiology.msu.edu/home_contacts.shtml)), respectively [8, 18, 33].

## Authors' contributions

ZT, MD, and YZ designed the experiments. MD and YZ performed most of the experiments. DT and XL made substantial contributions to the materials collection and data analysis. YZ wrote the manuscript. MD revised and edited the manuscript. ZT and MD gave the final approval of the manuscript. All authors read and approved the paper.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 19 April 2018 Accepted: 27 August 2018

Published online: 03 September 2018

## References

- Alvarez-Buylla ER, Liljegren SJ, Pelaz S, Gold SE, Burgeff C, Ditta GS, Vergara-Silva F, Yanofsky MF. MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. *Plant J*. 2000;24(4):457–66.
- Zhang H, Forde BG. An *Arabidopsis* MADS box gene that controls nutrient-induced changes in root architecture. *Science*. 1998;279(5349):407–9.
- Alvarez-Buylla ER, Pelaz S, Liljegren SJ, Gold SE, Burgeff C, Ditta GS, De Poupiana LR, Martínez-Castilla L, Yanofsky MF. An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proc Natl Acad Sci*. 2000;97(10):5328–33.
- Ng M, Yanofsky MF. Function and evolution of the plant MADS-box gene family. *Nat Rev Genet*. 2001;2(3):186–95.
- Kaufmann K, Melzer R, Theissen G. MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene*. 2005;347(2):183–98.
- Yang Y, Jack T. Defining subdomains of the K domain important for protein-protein interactions of plant MADS proteins. *Plant Molecular Biology*. 2004; 55(1):45–59.
- Nam J, Kim J, Lee S, An G, Ma H, Nei M. Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms. *Proc Natl Acad Sci*. 2004;101(7):1910–1915.
- Parenicová L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B. Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis* new openings to the MADS world. *Plant Cell*. 2003;15(7):1538–51.



9. Sanda SL, Amasino RM. Interaction of FLC and late-flowering mutations in *Arabidopsis thaliana*. *Mol Gen Genet MGG*. 1996;251(1):69–74.
10. Hartmann U, Höhmann S, Nettekheim K, Wisman E, Saedler H, Huijser P. Molecular cloning of SVP: a negative regulator of the floral transition in *Arabidopsis*. *Plant J*. 2000;21(4):351–60.
11. Moon J, Suh SS, Lee H, Choi KR, Hong CB, Paek NC, Kim SG, Lee I. The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in *Arabidopsis*. *Plant J*. 2003;35(5):613–23.
12. Mandel MA, Gustafson-Brown C, Savidge B, Yanofsky MF. Molecular characterization of the *Arabidopsis* floral homeotic gene APETALA1. *Nature*. 1992;360(6401):273–7.
13. Gu Q, Ferrándiz C, Yanofsky MF, Martienssen R. The FRUITFULL MADS-box gene mediates cell differentiation during *Arabidopsis* fruit development. *Development*. 1998;125(8):1509–17.
14. Liljegren SJ, Ditta GS, Eshed Y, Savidge B, Bowman JL, Yanofsky MF. SHATTERPROOF MADS-box genes control seed dispersal in *Arabidopsis*. *Nature*. 2000;404(6779):766–70.
15. Ma H. The ABCs of floral evolution. *Cell*. 2000;101(1):5–8.
16. Riechmann JL, Meyerowitz EM. MADS domain proteins in plant development. *Biological Chemistry*. 1997;378(10):1079–1101.
17. Nesi N, Debeaujon I, Jond C, Stewart AJ, Jenkins GI, Caboche M, Lepiniec L. 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*. 2002;14(10):2463–79.
18. Arora R, Agarwal P, Ray S, Singh AK, Singh VP, Tyagi AK, Kapoor S. MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. *BMC Genomics*. 2007;8(1):242.
19. Theissen G, Becker A, Di RA, Kanno A, Kim JT, Münster T, Winter KU, Saedler H. A short history of MADS-box genes in plants. *Plant Molecular Biology*. 2000;42(1):115–149.
20. Yanofsky MF, Ma H, Bowman JL, Drews GN, Feldmann KA, Meyerowitz EM. The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors. *Nature*. 1990;346(6279):35–9.
21. Weigel D, Meyerowitz EM. The ABCs of floral homeotic genes. *Cell*. 1994;78(2):203–9.
22. Masiero S, Colombo L, Grini PE, Schnittger A, Kater MM. The emerging importance of type I MADS box transcription factors for plant reproduction. *Plant Cell*. 2011;23(3):865–72.
23. Clark LG, Londoño X, Ruiz-Sanchez E. *Bamboo Taxonomy and Habitat*. Springer International Publishing, 2015.
24. Yuan JL, Yue JJ, Gu XP, Lin CS. Flowering of Woody bamboo in tissue culture systems. *Front Plant Sci*. 2017;8:1589.
25. Isagi Y, Shimada K, Kushima H, Tanaka N, Nagao A, Ishikawa T, Onodera H, Watanabe S. Clonal structure and flowering traits of a bamboo [*Phyllostachys pubescens* (mazel) Ohwi] stand grown from a simultaneous flowering as revealed by AFLP analysis. *Mol Ecol*. 2004;13(7):2017–21.
26. Wong KM. Flowering, fruiting and germination of the bamboo *Schizostachyum zollingeri* in Perlis. *Malays Forester*. 1981;44(4):453–63.
27. Campbell J. Bamboo flowering patterns: a global view with special reference to East Asia. *J Am Bamboo Soc*. 1985;6:17–35.
28. Troup RS, Troup RS. The silviculture of Indian trees. Published under the authority of his Majesty's secretary of state for India in council. *Cesk Pediatr*. 1921;12(1):1–12.
29. Keeley JE, Bond WJ. Mast flowering and Semelparity in bamboos: the bamboo fire cycle hypothesis. *Am Nat*. 1999;154(3):383.
30. Lin EP, Peng HZ, Jin QY, Deng MJ, Li T, Xiao XC, Hua XQ, Wang KH, Bian HW, Han N. Identification and characterization of two bamboo (*Phyllostachys praecox*) AP1 / SQUA-like MADS-box genes during floral transition. *Planta*. 2009;231(1):109–20.
31. Shih MC, Chou ML, Yue JJ, Hsu CT, Chang WJ, Ko SS, Liao DC, Huang YT, Chen JJ, Yuan JL. BeMADS1 is a key to delivery MADSs into nucleus in reproductive tissues- De novo characterization of *Bambusa edulis* transcriptome and study of MADS genes in bamboo floral development. *BMC Plant Biol*. 2014;14(1):179.
32. Gao J, Zhang Y, Zhang C, Qi F, Li X, Mu S, Peng Z. Characterization of the floral transcriptome of Moso bamboo (*Phyllostachys edulis*) at different flowering developmental stages by transcriptome sequencing and RNA-Seq analysis. *PLoS One*. 2014;9(6):e98910.
33. Wei B, Zhang R-Z, Guo J-J, Liu D-M, Li A-L, Fan R-C, Mao L, Zhang X-Q. Genome-wide analysis of the MADS-box gene family in *Brachypodium distachyon*. *PLoS One*. 2014;9(1):e84781.
34. Solovyev V, Kosarev P, Seledsov I, Vorobyev D. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol*. 2006;7(Suppl 1):1–12.
35. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29(7):644–52.
36. Milne I, Stephen G, Bayer M, Cock PJ, Pritchard L, Cardle L, Shaw PD, Marshall D. Using tablet for visual exploration of second-generation sequencing data. *Brief Bioinform*. 2013;14(2):193–202.
37. Sheen J. Signal transduction in maize and *Arabidopsis* mesophyll protoplasts. *Plant Physiol*. 2001;127(4):1466.
38. Katoh K, Kuma K, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res*. 2005;33(2):511–8.
39. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res*. 2015;43(D1):D257–60.
40. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res*. 2009;37(suppl\_2):W202–8.
41. Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci*. 1998;95(11):5857–64.
42. Letunic I, Doerks T, Bork P. SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res*. 2012;40(Database issue):302–5.
43. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S. PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res*. 2002;30(1):325.
44. Fan C, Ma J, Guo Q, Li X, Wang H, Lu M. Selection of reference genes for quantitative real-time PCR in bamboo (*Phyllostachys edulis*). *PLoS One*. 2013;8(2):e56573.
45. Zhao H, Lou Y, Sun H, Li L, Wang L, Dong L, Gao Z. Transcriptome and comparative gene expression analysis of *Phyllostachys edulis* in response to high light. *BMC Plant Biology*. 2016;16(1):34.
46. Zhao T, Ni Z, Dai Y, Yao Y, Nie X, Sun Q. Characterization and expression of 42 MADS-box genes in wheat (*Triticum aestivum* L.). *Mol Gen Genomics*. 2006;276(4):334–50.
47. Chen S, Songkumarn P, Liu J, Wang GL. A versatile zero background T-vector system for gene cloning and functional genomics. *Plant Physiol*. 2009;150(3):1111–21.
48. Clough SJ, Bent AF. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J Cell Mol Biol*. 1998;16(6):735–43.
49. Henschel K, Kofuji R, Hasebe M, Saedler H, Münster T, Theissen G. Two ancient classes of MIKC-type MADS-box genes are present in the moss *Physcomitrella patens*. *Mol Biol Evol*. 2002;19(6):801–14.
50. Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res*. 1999;27(2):573.
51. Folter SD, Angenent GC. Trans meets cis in MADS science. *Trends Plant Sci*. 2006;11(5):224–31.
52. Khan MR, Hu J, Ali GM. Reciprocal loss of CArG-boxes and auxin response elements drives expression divergence of MPF2-like MADS-box genes controlling Calyx inflation. *PLoS One*. 2012;7(8):e42781.
53. Gregis V, Sessa A, Dorca-Fornell C, Kater MM. The *Arabidopsis* floral meristem identity genes AP1, AGL24 and SVP directly repress class B and C floral homeotic genes. *Plant J*. 2010;60(4):626–37.
54. Torti S, Fornara F. AGL24 acts in concert with SOC1 and FUL during *Arabidopsis* floral transition. *Plant Signal Behav*. 2012;7(10):1251–4.
55. Gregis V, Sessa A, Colombo L, Kater MM. AGL24, SHORT VEGETATIVE PHASE, and APETALA1 Redundantly control AGAMOUS during early stages of flower development in *Arabidopsis*. *Plant Cell*. 2006;18(6):1373.
56. V G AS, D-F C, MM K. The *Arabidopsis* floral meristem identity genes AP1, AGL24 and SVP directly repress class B and C floral homeotic genes. *Plant J*. 2009;60(4):626–37.
57. Gramzow L, Weilandt L, Theissen G. MADS goes genomic in conifers: towards determining the ancestral set of MADS-box genes in seed plants. *Ann Bot*. 2014;114(7):1407.
58. Gramzow L, Ritz MS, Theissen G. On the origin of MADS-domain transcription factors. *Trends Genet*. 2010;26(4):149–53.
59. Leseberg CH, Li A, Kang H, Duvall M, Mao L. Genome-wide analysis of the MADS-box gene family in *Populus trichocarpa*. *Gene*. 2006;378:84–94.

60. Kapazoglou A, Engineer C, Drosou V, Kalloniati C, Tani E, Tsaballa A, Kouri ED, Ganopoulos I, Fletmetakis E, Tsafaris AS. The study of two barley type I-like MADS-box genes as potential targets of epigenetic regulation during seed development. *BMC Plant Biol.* 2012;12(1):166.
61. Peng Z, Lu Y, Li L, Zhao Q, Feng Q, Gao Z, Lu H, Hu T, Yao N, Liu K. The draft genome of the fast-growing non-timber forest species moso bamboo (*Phyllostachys heterocycla*). *Nat Genet.* 2013;45(4):456–61.
62. Messing J, Bharti AK, Karlowski WM, Gundlach H, Kim HR, Yu Y, Wei F, Fuks G, Soderlund CA, Mayer KF. Sequence composition and genome organization of maize. *Proc Natl Acad Sci U S A.* 2004;101(40):14349–54.
63. Jaillon O, Aury J-M, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, Aubourg S, Vitulo N, Jubin C. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature.* 2007;449(7161):463–7.
64. Nystedt B, Street NR, Wetterbom A, Zuccolo A, Lin Y-C, Scofield DG, Vezzi F, Delhomme N, Giacomello S, Alexeyenko A. The Norway spruce genome sequence and conifer genome evolution. *Nature.* 2013;497(7451):579–84.
65. Finnegan DJ. Eukaryotic transposable elements and genome evolution. *Trends Genet.* 1989;5:103–7.
66. Zhou M, Hu B, Zhu Y. Genome-wide characterization and evolution analysis of long terminal repeat retroelements in moso bamboo (*Phyllostachys edulis*). *Tree Genet Genomes.* 2017;13(2):43.
67. Ma J, SanMiguel P, Lai J, Messing J, Bennetzen JL. DNA rearrangement in orthologous orp regions of the maize, rice and sorghum genomes. *Genetics.* 2005;170(3):1209–20.
68. SanMiguel P, Tikhonov A, Jin Y-K, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z. Nested retrotransposons in the intergenic regions of the maize genome. *Science.* 1996;274(5288):765–8.
69. Bowman JL, Drews GN, Meyerowitz EM. Expression of the Arabidopsis floral homeotic gene AGAMOUS is restricted to specific cell types late in flower development. *Plant Cell.* 1991;3(8):749–58.
70. Mizukami Y, Ma H. Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity. *Cell.* 1992;71(1):119–31.
71. Theißen G, Saedler H. Plant biology: floral quartets. *Nature.* 2001;409(6819):469–71.
72. Theißen G, Kim JT, Saedler H. Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. *J Mol Evol.* 1996;43(5):484–516.
73. Gao XC, Liang WQ, Yin CS, Ji SM, Wang HM, Xiao S, Guo CC, Kong HZ, Xue HW, Zhang DB. The SEPALLATA-like gene OsMADS34 is required for rice inflorescence and spikelet development. *Plant Physiol.* 2010;153(2):728–40.
74. Liu C, Xi W, Shen L, Tan C, Yu H. Regulation of floral patterning by flowering time genes. *Dev Cell.* 2009;16(5):711–22.
75. Yu H, Xu Y, Tan EL, Kumar PP. AGAMOUS-LIKE 24, a dosage-dependent mediator of the flowering signals. *Proc Natl Acad Sci U S A.* 2002;99(25):16336–41.
76. Liu C, Chen H, Er HL, Soo HM, Kumar PP, Han JH, Liou YC, Yu H. Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. *Development.* 2008;135(8):1481.
77. Lee JH, Yoo SJ, Park SH, Hwang I, Lee JS, Ahn JH. Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. *Genes Dev.* 2007;21(4):397.
78. Liu C, Zhou J, Brachadri K, Yalovsky S, Ito T, Yu H. Specification of Arabidopsis floral meristem identity by repression of flowering time genes. *Development.* 2007;134(10):1901.
79. Lee JH, Park SH, Ji HA. Functional conservation and diversification between rice OsMADS22/OsMADS55 and Arabidopsis SVP proteins. *Plant Sci.* 2012;185-186:97–104.
80. Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF. B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature.* 2000;405(6783):200.
81. Peña L, Martíntrillo M, Juárez J, Pina JA, Navarro L, Martínezpater JM. Constitutive expression of Arabidopsis LEAFY or APETALA1 genes in citrus reduces their generation time. *Nat Biotechnol.* 2001;19(3):263–7.
82. Goto K. Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature.* 2001;409(6819):525.
83. Kramer EM, Jaramillo MA, Di SV. Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms. *Genetics.* 2004;166(2):1011–23.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

