Plant Biotechnology Journal

Plant Biotechnology Journal (2015), pp. 1–15 doi: 10.1111/pbi.12383

Transcriptome-wide analysis of the MADS-box gene family in the orchid *Erycina pusilla*

Choun-Sea Lin[1,*], Chen-Tran Hsu[1], De-Chih Liao[1], Wan-Jung Chang[1], Ming-Lun Chou[2], Yao-Ting Huang[3], Jeremy J. W. Chen[4], Swee-Suak Ko[1,5], Ming-Tsair Chan[1,5] and Ming-Che Shih[1,*]

[1]*Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan*

[2]*Department of Life Sciences, Tzu Chi University, Hualien, Taiwan*

[3]*Department of Computer Science and Information Engineering, National Chung Cheng University, Chia-yi, Taiwan*

[4]*Institute of Biomedical Sciences, National Chung Hsing University, Taichung, Taiwan*

[5]*Biotechnology Center in Southern Taiwan, Academia Sinica, Tainan, Taiwan*

Received 21 January 2015;
revised 5 March 2015;
accepted 18 March 2015.
**Correspondence* (Tel +886-2-27872094;
fax +886-2-2651-5600;
email cslin99@gate.sinica.edu.tw;
Tel +2-27872002;
fax +886-2-2651-5600;
email mcshih@gate.sinica.edu.tw)

Keywords: flowering, bacteria artificial chromosome, subcellular localization, next-generation sequencing.

Summary

Orchids exhibit a range of unique flower shapes and are a valuable ornamental crop. MADS-box transcription factors are key regulatory components in flower initiation and development. Changing the flower shape and flowering time can increase the value of the orchid in the ornamental horticulture industry. In this study, 28 MADS-box genes were identified from the transcriptome database of the model orchid *Erycina pusilla*. The full-length genomic sequences of these MADS-box genes were obtained from BAC clones. Of these, 27 were MIKC-type EpMADS (two truncated forms) and one was a type I EpMADS. Eleven *EpMADS* genes contained introns longer than 10 kb. Phylogenetic analysis classified the 24 MIKC^{c} genes into nine subfamilies. Three specific protein motifs, AG, FUL and SVP, were identified and used to classify three subfamilies. The expression profile of each *EpMADS* gene correlated with its putative function. The phylogenetic analysis was highly correlated with the protein domain identification and gene expression results. Spatial expression of *EpMADS6*, *EpMADS12* and *EpMADS15* was strongly detected in the inflorescence meristem, floral bud and seed via *in situ* hybridization. The subcellular localization of the 28 EpMADS proteins was also investigated. Although EpMADS27 lacks a complete MADS-box domain, EpMADS27-YFP was localized in the nucleus. This characterization of the orchid MADS-box family genes provides useful information for both orchid breeding and studies of flowering and evolution.

Introduction

The Orchidaceae is one of the largest families of flowering plants (Aceto and Gaudio, 2011). Orchids exhibit a myriad of unique floral shapes, colours and fragrances and are an important ornamental crop (Chandler and Sanchez, 2012). Orchid flower initiation and development has been studied at the molecular level (as reviewed by Hsiao *et al.*, 2011) to potentially enable future direct breeding.

Plant transformation represents one direction for ornamental plant breeding. As ornamental plants are not for human consumption, the marketplace has generally been found to be accepting of genetically modified varieties (Chandler and Sanchez, 2012; Thiruvengadam *et al.*, 2012). Indeed, the transgenic approach is currently used for some orchid breeding (Ding *et al.*, 2013; Thiruvengadam *et al.*, 2012). The major processes of life depend on differential gene expression, which is largely controlled and regulated by the activity of transcription factors. The MADS-box family of transcription factors is involved in the regulation of developmental processes and environmental responses in eukaryotic organisms (Ng and Yanofsky, 2001; Theißen and Saedler, 2001; Theißen *et al.*, 2000). MADS-box proteins contain a highly conserved approximately 60-amino-acid-long sequence in the N-terminal region (M-domain) termed the MADS-box (Theißen *et al.*, 2000). The MADS-box gene family can be divided into two lineages, type I and type II, based on the exon/intron and domain structures, rates of evolution, developmental function and degree of functional redundancy (Bemer *et al.*, 2010; Gramzow *et al.*, 2010; Nam *et al.*, 2004; Smaczniak *et al.*, 2012). Plant type I MADS-domain genes can be further subdivided into three groups: Mα, Mβ and Mγ (Bemer *et al.*, 2010; Parenicová *et al.*, 2003). Type II MADS-box proteins are commonly referred to as MIKC-type proteins based on their domain structure, which comprises the highly conserved MADS-domain (M), a moderately conserved intervening (I) domain, a well-conserved keratin-like (K) domain and a highly divergent carboxy terminal (C) domain (Kaufmann *et al.*, 2005b; Parenicová *et al.*, 2003; Smaczniak *et al.*, 2012; Verelst *et al.*, 2007). The MIKC-type MADS-box genes have been further divided into two subtypes, MIKC^{C} and MIKC*, based on sequence divergence including and after the I-domain (Henschel *et al.*, 2002). The functional diversification of MIKC^{C} genes accounts for the developmental and morphological novelties in the sporophytic generation of seed plants (particularly the floral organs and fruits of angiosperms), which can be explained in terms of the ABCDE model (Arora *et al.*, 2007; Kaufmann *et al.*, 2005b; Liu and Mara, 2010; Theißen *et al.*, 2000). The combination of

Please cite this article as: Lin, C.-S., Hsu, C.-T., Liao, D.-C., Chang, W.-J., Chou, M.-L., Huang, Y.-T., Chen, J.J.W., Ko, S.-S., Chan, M.-T. and Shih, M.-C. (2015) Transcriptome-wide analysis of the MADS-box gene family in the orchid *Erycina pusilla*. *Plant Biotechnol. J.*, doi: 10.1111/pbi.12383

functions in the floral meristem cells determines the later floral organ identity: sepal (A and E), petal (A, B, and E), stamen (B, C, and E), carpel (C and E) and ovule (D and E) (Kaufmann *et al.*, 2005b; Ma and dePamphilis, 2000; Weigel and Meyerowitz, 1994). Therefore, most studies of orchid MADS-box genes have focused on these ABCDE genes and employed homology searches to identify orchid MADS-box genes (Aceto and Gaudio, 2011; Hsiao *et al.*, 2011). The identification and functional characterization of B-, C/D- and E-class MADS-box genes in the specification of orchid floral organ development have been reported for only a few species, such as *Phalaenopsis*, *Oncidium* and *Dendrobium* (Aceto and Gaudio, 2011; Acri-Nunes-Miranda and Mondragon-Palomino, 2014; Chang *et al.*, 2009; Chen *et al.*, 2007, 2012; Hsu *et al.*, 2010; Mondragon-Palomino, 2013; Tsai *et al.*, 2008; Xu *et al.*, 2006). Gene duplications within either functional classes or phylogenetic clades have also been analysed (Acri-Nunes-Miranda and Mondragon-Palomino, 2014; Mondragon-Palomino, 2013). The Homeotic Orchid Tepal (HOT) model has been proposed to explain how orchid MADS-box genes regulate floral organ development. The HOT model surmises that divergent genes in the B-class combine with other classes of MADS-box genes to regulate the complexity of sepal, petal and lip identity (Pan *et al.*, 2011) and that the relative quantities of the highly expressed B-class genes distinguish orchid tepal development (a revision of the 'orchid code' model by Mondragon-Palomino and Theißen, 2011). Accordingly, it is generally hypothesized that higher-order protein complexes formed by multiple homeotic MADS-box proteins determine orchid floral organ identity.

Transformation and virus-induced gene silencing (VIGS) can be used to control the expression of MADS genes or the expression of a chimeric repressor of MADS to manipulate the flowering time and flower morphology of ornamental crops (Chandler and Sanchez, 2012; Ding *et al.*, 2013; Sage-Ono *et al.*, 2013; Tanaka *et al.*, 2013; Thiruvengadam *et al.*, 2012; Yu *et al.*, 1999). In *Cyclamen persicum*, expression of a chimeric repressor of CpAG1 could convert stamens into petals, resulting in an increase in petal number from 5 to 10 (Tanaka *et al.*, 2013). However, expression of a chimeric repressor of CpAG2, only caused incomplete formation of stamens and carples, even though CpAG1 and CpAG2 share 90% amino acid identity (Tanaka *et al.*, 2013). Therefore, to understand the global role of MADS-box proteins in growth and development and to obtain information applicable to crop production, it is pertinent to examine the entire MADS-box gene family. *Erycina pusilla* is attractive for use as a model orchid for functional genomic and genetic analyses (Pan *et al.*, 2012) because of its low diploid chromosome number (2n, $n = 6$), relatively short juvenile phase and, most importantly, its ability to complete its life cycle *in vitro* (Chiu *et al.*, 2011). *Erycina pusilla* miRNA BAC libraries (Lin *et al.*, 2013) have been constructed, and the transcriptome (Chou *et al.*, 2013), tissue culture system (Chiu *et al.*, 2011) and transformation protocol (Lee *et al.*, 2015) have been established. In this study, 28 MADS-box genes were identified from the *E. pusilla* transcriptome. Molecular, bioinformatics and cell biology tools were then used to investigate these 28 genes to obtain a global view of the MADS-box gene family in Orchidaceae.

Results and discussion

Identification of *E. pusilla* MADS-box domain genes in the transcriptome database

A search of the *E. pusilla* transcriptome database (Chou *et al.*, 2013) revealed several contigs (DNA fragments) with sequences similar to the MADS-box genes identified in other plant species. Specific primers were designed based on these contig sequences and used to screen the BAC library (Pan *et al.*, 2012). In total, 28 *E. pusilla* MADS-box genes (*EpMADS1 to 28*) were isolated. After BLASTP analysis, these 28 EpMADSs were classified as type I or type II. The number of genes identified was lower than that identified in other plant species (Table 1). We detected 24 MIKC^c-type genes (*EpMADS1* to *24*), whereas other monocots contain 33 or 39 MIKC^c-type genes (sorghum and maize/rice, Table 1). In addition, compared with the number of MADS-box contigs identified here in *E. pusilla*, recent transcriptome-based studies have identified more in other orchids, for example 158 in *Cymbidium ensifolium* (Li *et al.*, 2013), 73 in *Cymbidium sinense* (Zhang *et al.*, 2013), 56 in *Ophrys* (Sedeek *et al.*, 2013) and more than 50 in *Phalaenopsis aphrodite* (Su *et al.*, 2011). However, most of these contigs were isoforms. In the putative E-class unigene comp55336_c0 in *C. ensifolium*, 21 contigs have identical sequences (Li *et al.*, 2013). In published orchid databases, many MADS-box contigs are only partial sequences. The *EpMADS* genes presented in this report are all full length and were confirmed by genomic DNA sequencing. For this reason, the number of genes may be smaller than those found in other orchid transcriptome databases, which include fragments of MADS-box genes. In addition, an initial search of the transcriptome may underestimate the number of *EpMADS* genes because MADS-box proteins are transcription factors and are frequently expressed at low levels, particularly in the case of the type I genes (Arora *et al.*, 2007). Because the whole-genome sequence of *E. pusilla* is not available, additional genes containing the MADS-box domain may be identified as more transcriptome and genome sequence information becomes available. Deeper transcriptome sequencing of materials from various developmental stages may identify additional MADS-box genes, particularly type I and MIKC^c type.

Using transcriptome and BAC sequence data, the full-length cDNA sequences and intron/exon positions of the 28 *EpMADS* genes were obtained (Figure 1). The mRNA activity and splice accuracy of the predicted genes were confirmed by RT-PCR. The gene structures of the MIKC^c-type genes were found to contain 7–10 exons, with the exception of *EpMADS17*, which contains 5. Type II MADS-box genes in rice, maize, sorghum, grape and *Arabidopsis* also contain multiple introns (Arora *et al.*, 2007; Díaz-Riquelme *et al.*, 2009; Parenicová *et al.*, 2003; Zhao *et al.*, 2011). We identified two truncated genes in the transcriptome, *EpMADS26* and *EpMADS27* that contain only an M- or I-domain, respectively. The missing M-, I-, K-, and C-domain sequences were not identified in the *EpMADS26* and *EpMADS27* BAC clones, possibly due to a long intron creating partial sequences in both the transcriptome and BAC libraries. Identification of the *EpMADS* genomic sequences will provide useful information for promoter and evolutionary studies.

In contrast to *Arabidopsis* (Kofuji *et al.*, 2003), nearly all of the *E. pusilla* MADS-box genes have introns of greater than 10 kb (Figure 1). Large introns have also been observed in *Oncidium* and *Orchis italica* MADS-box and other flowering-related orchid genes (Hsu *et al.*, 2011; Salemme *et al.*, 2013). Many MADS-box genes containing large first and second introns have been identified in grape (Díaz-Riquelme *et al.*, 2009). However, large introns are distributed throughout the *EpMADS* genes, suggesting that it will be difficult to sequence, assemble and predict genes from the complex genomes of orchids, including *E. pusilla*.

The intron splice sites are highly conserved in most MIKC^c-type MADS-box genes in *Arabidopsis*, grape and cucumber (Arora

Table 1 Categorization of MADS-box gene families in nine plant genomes

Species	*E. pusilla*	Tomato	Poplar	Grape	Maize	Sorghum	Rice	Arabidopsis	Cucumber
Total type II genes	27	36	64	38	43	35	43	45	33
MIKC^{C†}	26 (9)	36 (12)	55 (12)	38 (13)	39 (9)	33 (10)	39 (12)	39 (12)	30 (10)
SOC	2	6	4	3	3	3	3	6	5
AGL6	3	1	1	1	3	2	2	2	2
SEP	4	5	5	3	3	5	5	4	4
AP1-FUL	3	5	6	3	5	4	4	4	3
AP3-PI	4	4	3	3	6	3	4	2	3
SVP	2	2	8	5	3	4	3	2	4
AG	4	4	4	3	4	4	5	4	1
CFO	1	–	–	–	–	–	1	–	–
B sister	1	1	3	2	6	2	3	2	–
AGL12	–	1	4	1	3	3	2	1	–
AGL15	–	1	2	2	–	–	1	2	2
AGL17	–	2	5	2	–	1	5	4	3
FLC	–	3	6	2	–	–	–	6	–
TM8	–	1	–	1	–	–	–	–	2
Unclassified	2	–	4	–	3	2	–	–	
MIKC*	1		9	–	4	2	4	8	3
Total type I genes‡	1	–	41	–	32	30	32	61	10
Mα	–	–	23	–	27	26	13	25	5
Mβ	–	–	12	–	3	2	9	20	2
Mγ	1	–	6	–	2	2	10	16	3
Total*	28	36	105	38	75	65	75	107	43

*Total number of identified type I and type II MADS genes in a given species.

†The MIKC^{C}-type II gene family can be divided into 15 subfamilies. The number in parentheses is the number of subfamilies present in that species.

‡Type I classification was defined by *Arabidopsis* type I genes.

et al., 2007; Díaz-Riquelme *et al.*, 2009; Hu and Liu, 2012; Parenicová *et al.*, 2003; Zhao *et al.*, 2011). In *EpMADS*, the lengths of exons 1, 2, 3, 4, 5 and 6 are conserved with respect to *Arabidopsis* and grape (Arora *et al.*, 2007; Díaz-Riquelme *et al.*, 2009; Parenicová *et al.*, 2003). Likewise, the first ~60 amino acids (M-domain) tend to be in the first exon. The second exon typically includes the last 15 amino acids of the M-domain and the I-domain. In addition, the K-domain is mainly encoded by exons 3, 4, 5 and 6 (Arora *et al.*, 2007; Díaz-Riquelme *et al.*, 2009; Parenicová *et al.*, 2003). Thus, some gene structures are conserved between orchids and other model crops.

Phylogenetic analysis of EpMADS

To determine the evolutionary relationships and classify the *E. pusilla* MIKC-type MADS-box proteins, a neighbour-joining phylogenetic tree was constructed. The 24 EpMADS proteins were classified into 9 MIKC^{C} subfamilies (Table 1; Figures 2a and S1). Every species analysed thus far contains members of the following MADS-box domain subfamilies: SUPPRESSOR OF OVEREXPRESSION OF CO (SOC), AGAMOUS-like 6 (AGL6), SEPALLATA (SEP), APETALA1/FRUITFULL (AP1-FUL), APETALA3/PISTILLATA (AP3-PI), SHORT VEGETATIVE PHASE (SVP) and AGAMOUS (AG). FLOWERING LOCUS C (FLC), AGL12, AGL15, AGL17 and tomato MADS-box gene 8 (TM8) subfamilies were absent in the *E. pusilla* transcriptome. These subfamilies are also absent in other species (Table 1; Figure S1).

The functions of some of these subfamilies may explain their absence in the tropical orchid *E. pusilla*. FLC proteins act as floral repressors regulated by vernalization and function in autonomous pathways that control the transition to flowering in *Arabidopsis* (Michaels and Amasino, 1999; Rouse *et al.*, 2002). According to Vaz *et al.* (2004), the optimal temperature for promotion of flowering in *E. pusilla* is 27 °C, and thus, vernalization does not affect its flowering. Therefore, like rice, *E. pusilla* does not need an active FLC protein to complete its life cycle (Arora *et al.*, 2007).

In *Arabidopsis*, AGL12 responds to root meristem cell proliferation and floral transition (Tapia-López *et al.*, 2008). Thus far, only cucumber and *E. pusilla* have been shown not to contain members of this gene family. It is unclear whether this limited absence is due to incomplete data from unsequenced genomes or whether the genes are truly absent in *E. pusilla*. In the gymnosperm *Gnetum gnemon*, *GGM10* appears to be a member of a sister group of the *AGL12*-like genes from angiosperms (Becker *et al.*, 2000). However, *GGM10* expression could not be detected by northern blot hybridization of RNA from leaves or male or female reproductive cones (Becker *et al.*, 2000). *AGL12*-like genes may be expressed in *E. pusilla* at levels too low to detect. Therefore, deeper transcriptome sequencing or genomic sequencing is required to further determine whether *AGL12*-like genes are present in *E. pusilla*.

In the *E. pusilla* transcriptome, we detected only one gene, EpMADS17, from the CFO subfamily (Figure 2a). Although CFO is a typical MIKC^{C}-type protein (Arora *et al.*, 2007; Nam *et al.*, 2004), its phylogenetic relationships remain unclear. CFO1 and its orthologues in grasses and *Phoenix dactylifera* have been classified as a monocot-specific subfamily thought to play a pivotal role in maintaining floral organ identity via negative regulation of the expression of another floral organ identity gene, *DROOPING LEAF* (*DL*) (Sang *et al.*, 2012).

SOC1 is a floral integrator that down-regulates the activity of other regulators, such as the floral meristem identity gene LEAFY (LFY), to determine the formation of floral meristems (Ding *et al.*,

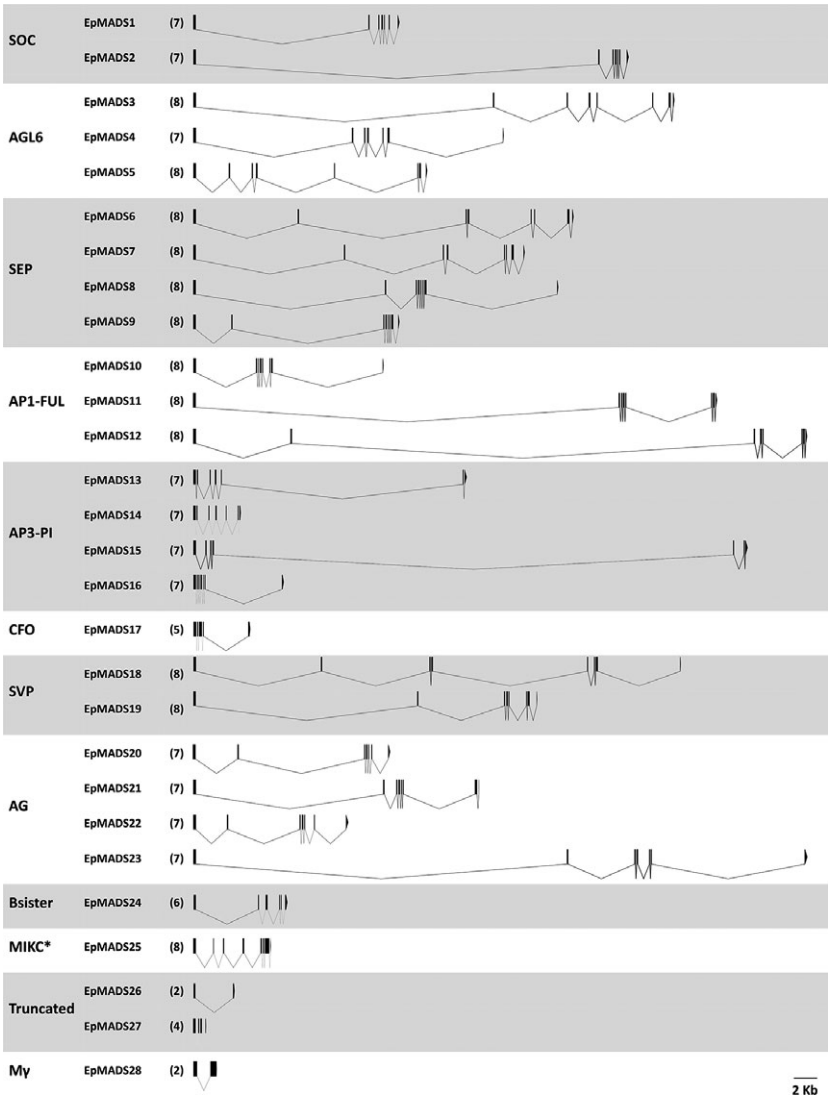

Figure 1 Exon/intron structures of *EpMADS* genes. The lines indicate introns, and the rectangles indicate exons. The numbers in parentheses represent the number of exons. EpMADS genes were divided into 12 subfamilies according to Parenicová *et al.* (2003) and numbered consecutively after this ordering.

2013; Lee and Lee, 2010; Wang *et al.*, 2009). Our phylogenetic analysis assigned EpMADS1 and EpMADS2 to the SOC subfamily. EpMADS1 is closely related to FOREVER YOUNG FLOWER (FVF) from *Oncidium* 'Gower Ramsey' and PeSOC1 from *Phalaenopsis equestris* (Figure 2b).

A large number of MADS-box genes have been cloned from orchid species and functionally characterized (reviewed in Acri-Nunes-Miranda and Mondragon-Palomino, 2014). Most of these studies have focused on the similarity of domain structures and the expression of genes homologous to *Arabidopsis* floral organ identity genes. Because the relationship between floral organ identity and MADS-box genes has been established by the ABCDE model, we discuss below the phylogeny of the EpMADS and other known orchid A/B/C/D/E-class MADS-box genes in this context.

The A-class group belongs to the AP1-FUL subfamily and function in the specification of sepal and petal identity (Coen and Meyerowitz, 1991). Phylogenetic analysis assigned EpMADS10/11/12 to the A-class group (Figure 2b). Two B-class MADS-box genes, *AP3* and *PI*, are required to specify petal and stamen identity (Coen and Meyerowitz, 1991). EpMADS13/14/15/16 were assigned to the B-class (AP3-PI) group (Figure 2b). The C-class group, including *Arabidopsis* AG and *Antirrhinum* PLENA (PLE), regulates the development of carpels and, together with the B-class group, stamens (Coen and Meyerowitz, 1991). EpMADS20/21/22 were assigned to the C-class (AG) group (Figure 2b). The D-class group is primarily involved in regulating ovule identity and includes *Arabidopsis* SEEDSTICK (STK, AG subfamily) and its petunia orthologues, petunia floral binding proteins 7 and 11 (FBP7/FBP11) (Colombo *et al.*, 1995; Pinyopich *et al.*, 2003). EpMADS23 was assigned to the D-class (STK) group (Figure 2b). The E-class group belongs to the SEP-like and AGL6 subfamily (AGL2/3/4; Malcomber and Kellogg, 2005; Theißen *et al.*, 2000; Theißen and Saedler, 2001; Zahn *et al.*, 2005). Three genes, EpMADS3/4/5, were assigned to the E-class (AGL6). EpMADS3 and EpMADS5 are closely related to MADS7 and MADS1 derived from *Oncidium* 'Gower Ramsey' (Figure 2b). The other four genes (EpMADS6/7/8/9) were assigned to the E-class (SEP) group, 6 and 7 in the LOF lineage and 8 and 9 in the SEP lineage (Figure 2b).

The B sister group of MADS-box genes are close relatives of the main B-class group. In eudicots, B sister MADS-box genes are involved in cell differentiation during ovule and seed development (Erdmann *et al.*, 2010; Kaufmann *et al.*, 2005a; Prasad *et al.*,

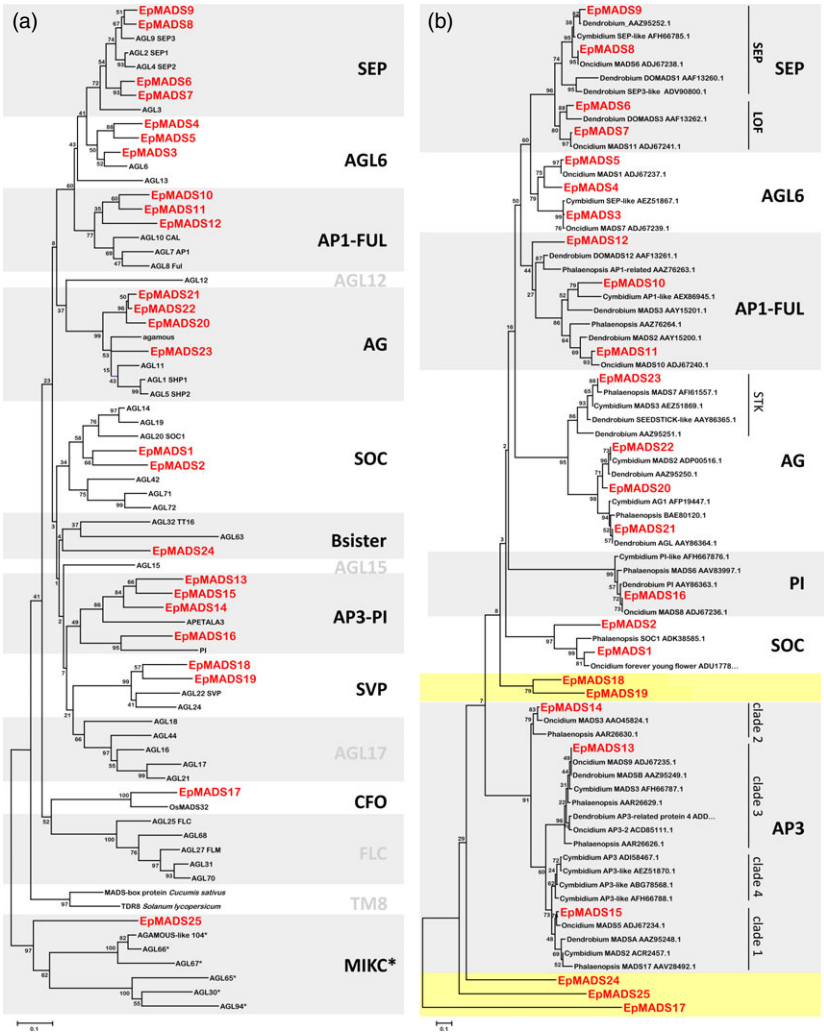

Figure 2 Phylogenetic analysis of the EpMADS subfamily. The phylogenetic tree was generated by the neighbour-joining (NJ) algorithm using Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0 (Tamura *et al.*, 2007). The numbers on major branches indicate bootstrap percentages for 1000 replicate analyses.
(a) Phylogenetic tree of *Arabidopsis* and *E. pusilla* MIKC proteins. Families that occur in *Arabidopsis* but for which no *E. pusilla* orthologues were identified are printed in light grey.
(b) Phylogenetic tree of the MADS-box proteins from the orchids *Phalaenopsis*, *Dendrobium*, *Cymbidium*, *Oncidium* and *E. pusilla*. Based on this figure and accompanying analyses, the *EpMADS* genes were divided into 10 subfamilies. These subfamilies are labelled at the right margin.

2010). EpMADS24 was assigned to the B sister group and is closely related to AGL32 and AGL63 derived from *Arabidopsis* (Figure 2a).

Domain identification of EpMADS

The overall protein structures of the predicted EpMADS proteins were analysed (Figures 3 and S2). The EpMADS-boxes have an expected length of 60 amino acids, are located in the N-terminal region and consist of motifs 1 and 3. The M-domains were found to be more highly conserved than the I-, K- or C-domains in the EpMADS. The I-domains are moderately conserved and vary in length, and the K-domains comprise motifs 2, 5 and 6 and are well conserved. The C-domains in the C-terminal region, however, were found to be highly divergent among the analysed EpMADS proteins.

EpMADS25 was the only MIKC*-type MADS-box protein identified in *E. pusilla* transcriptome. Other plant species have more than 1 MIKC*-type MADS-box protein, and it would be interesting to further investigate the number of MIKC*-type MADS-box proteins in *E. pusilla*.

The C-domain is the most diverse among the EpMADS proteins. This MADS-box domain responds to transcriptional activation and the formation of higher-order protein complexes (reviewed by Díaz-Riquelme *et al.*, 2009). The AGL6 and SEP subfamilies contain motif 7, whereas proteins in the AG family contain motif 8 (Figures 3 and S2). The motif 8 regions from *E. pusilla* (EpMADS20 to 23), four orchid species (eight proteins) and *Arabidopsis* (AG, AGL11, SHATTERPROOF 1 (SHP1), and SHP2) were aligned. Motif 8 contains three conserved stretches of sequence, A, B and C (Figure 4). In orchids, these proteins can be divided to two subgroups: Group I containing EpMADS20/21/22 and Group II containing EpMADS23. Group I EpMADS are most similar to *Arabidopsis* proteins in the B- and C-domains. The four orchid homologues in Group II (from *Dendrobium*, *Phalaenopsis* and *E. pusilla*) are more closely related to each other than to the proteins from *Arabidopsis* or EpMADS20/21/22.

In the SVP family, motif 13 is conserved in EpMADS18 and EpMADS19 (Figure 4). This motif is present in the transcriptional repressor SVP and the transcriptional enhancer AGL24, but no specific function of this motif has been assigned (Hartmann *et al.*, 2000; Wu *et al.*, 2012; Yu *et al.*, 2002).

Motif 14 is considered the paleoAP1 motif and is typical of the paleoAP1-like clade of the C-terminal region of AP1-like proteins (Figure 4). Among EpMADS10/11/12 in the AP1-FUL subfamily, motif 14 exhibits a unique pattern of conservation (Figure 4). Whereas the M- and I-domains are highly conserved within the AP1-FUL subfamily, the conserved motif 14 is present only in proteins most closely related overall to AtFUL. EpMADS11 and

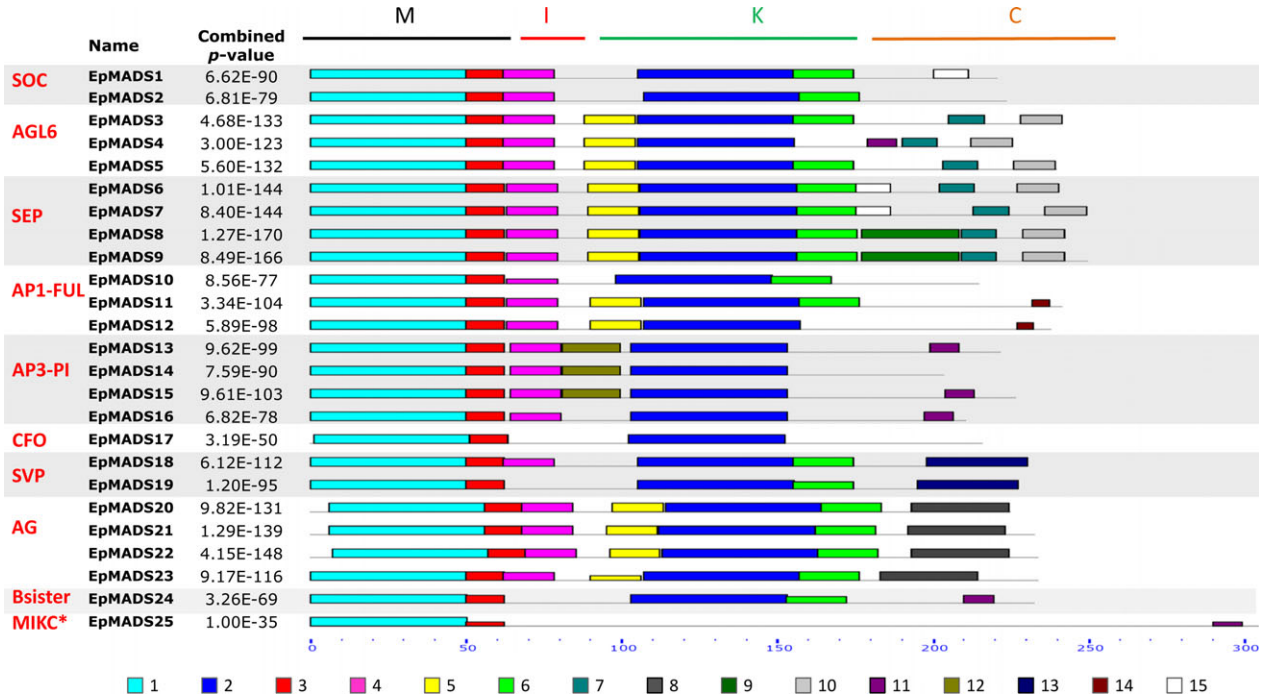

Figure 3 Organization of putative motifs in EpMADS. Putative motifs in EpMADS proteins identified by MEME (Bailey *et al.*, 2009). The numbered boxes represent different putative motifs. The expected values for all motifs were calculated using MEME and are displayed after each gene name.

EpMADS12 share 53% and 45% overall identity, respectively, with *Arabidopsis* FUL. Within the monocots, this domain is highly conserved in the AP1-like protein, but its function remains unknown. Three FUL-like orchid proteins, EpMADS10, AEX86945 and AAY15201, do not contain this motif. No orchid protein closely resembles AP1, CAULIFLOWER (CAL) or AtAGL13. Whereas the three *E. pusilla* FUL-like proteins are more similar to each other than to any *Arabidopsis* protein (FUL clade, Figure 2), the presence of motif 14 may indicate which *E. pusilla* protein performs the FUL function in orchids.

Gene expression of *EpMADS* in various tissues

The expression patterns of the 28 MADS-box genes (including the two truncated forms) in *E. pusilla* were analysed by quantitative RT-PCR using total RNA isolated from root, leaf, peduncle, flower, fruit and protocorm (Figure 5). The *SOC* subfamily members

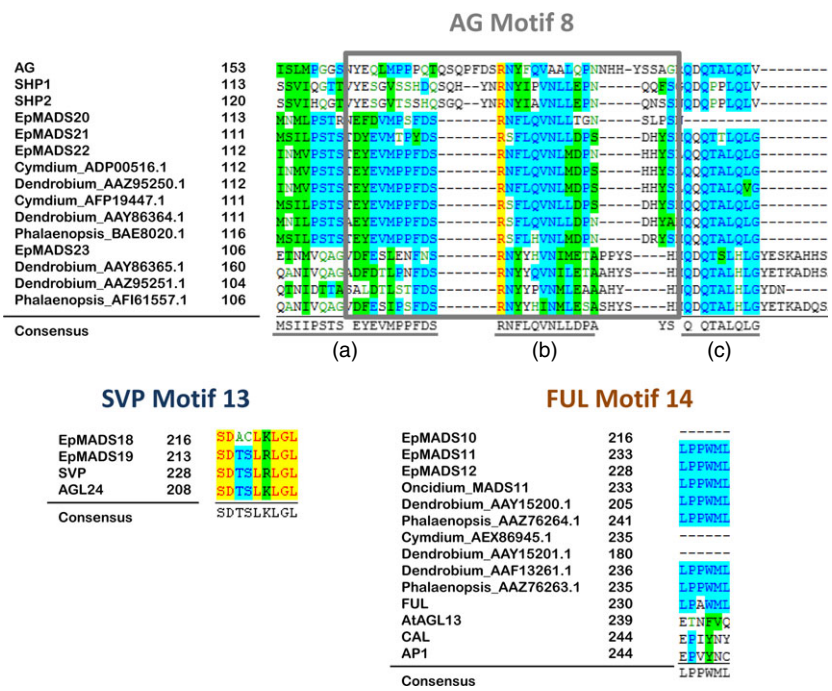

Figure 4 Unique C-terminal motifs. Amino acid CLUSTAL alignments were constructed for motifs present in different subfamilies of the MADS-box proteins of *Arabidopsis* and several orchid species.

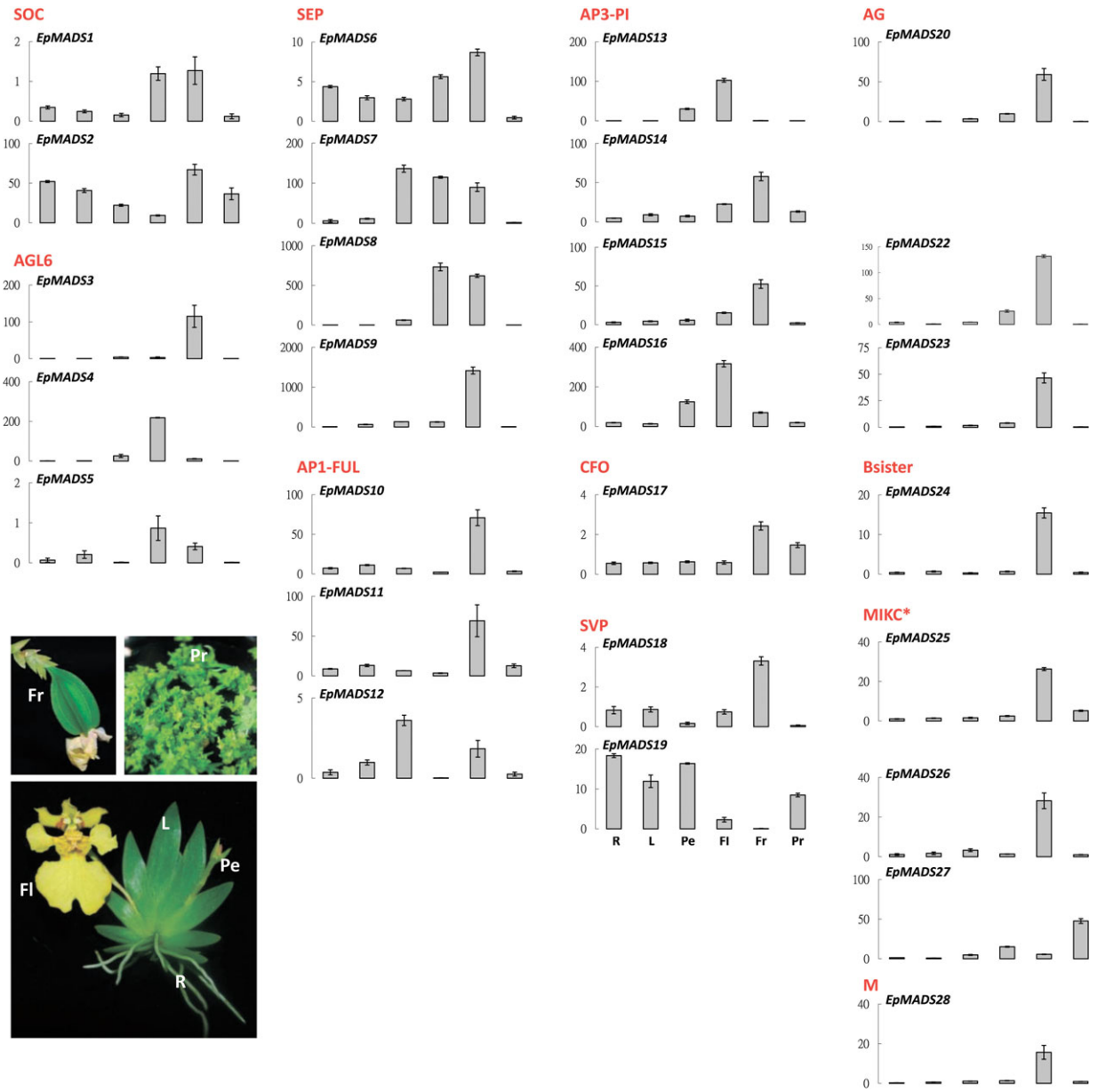

Figure 5 Tissue-specific expression patterns of *EpMADS* genes. RNA was extracted from five different tissues of *E. pusilla* and used for cDNA synthesis: R, root; L, leaf, Pe, peduncle; Fl, flower; Fr, fruits; Pr, protocorm. Real-time PCR was performed using the primers listed in Table S1.

EpMADS1 and *EpMADS2* were expressed in all tissues (Figure 5), similar to Arabidopsis *SOC1* and *AGL42* (Hepworth *et al.*, 2002) and rice *OsMADS56* and *OsMADS50* (Arora *et al.*, 2007). However, not all genes in the *SOC* subfamily are constitutively expressed. *AGL19* and *AGL14* are only expressed in the root, and three *SOC* subfamily genes were strongly detected in the bud, shoot and leaf in grape (Díaz-Riquelme *et al.*, 2009; Parenicová *et al.*, 2003).

The expression patterns of the *E. pusilla AGL6* subfamily genes differed. *EpMADS3* was expressed only in fruits; *EpMADS4* and *EpMADS5* were highly expressed in the flower (Figure 5). Similar expression patterns were observed in *Oncidium*: *OMADS6* (the homologue of *EpMADS3*) and *OMADS1* (the homologue of *EpMADS5*) were expressed in the reproductive tissues (Chang *et al.*, 2009; Hsu *et al.*, 2003). *SEP* subfamily genes, *EpMADS6/7/8/9,* were expressed in reproductive tissues with expression pattern similar to that observed for other orchid *SEP* orthologues (Acri-Nunes-Miranda and Mondragon-Palomino, 2014; Chang *et al.*, 2009; Pan *et al.*, 2014). This expression pattern is consistent with the function of *SEP* as an E-class gene (Kaufmann *et al.*, 2005b).

Expression of the *E. pusilla AP1-FUL* subfamily genes *EpMADS10* and *EpMADS11* was more dominant in fruits (Figure 5). High expression of *EpMADS12* was detected in the peduncle (Figure 5). In *Phalaenopsis* 'Athens', *PhaMADS1* and *PhaMADS2* are transcribed at a uniformly low level in the perianth and gynostemium and at a high level in the ovary before pollination (Acri-Nunes-Miranda and Mondragon-Palomino, 2014).

The majority of genes in the *E. pusilla AP3-PI* subfamily were highly expressed in reproductive tissues. *EpMADS13* and

EpMADS16 were expressed predominantly in the flower and *EpMADS14* and *EpMADS15* in fruits (Figure 5). Orchid *AP3*-like genes are grouped into four ancient clades, within which genes are expressed in the same organs of the perianth. Genes classified into clade 1 (*PeMADS2*-like genes) and clade 2 (*OMADS3*-like genes) are expressed in all tepals (Hsu and Yang, 2002; Tsai *et al.*, 2004; Xu *et al.*, 2006), whereas genes classified into clade 3 (*PeMADS3*-like genes) are only expressed in the inner tepal (Kim *et al.*, 2007; Tsai *et al.*, 2004; Xu *et al.*, 2006), and those in clade 4 (*PeMADS4*-like genes) are exclusively expressed in the labellum (Tsai *et al.*, 2004). This suggests a simple combinatorial model, or orchid code, for floral organ identity specification (Mondragon-Palomino and Theißen, 2008, 2009).

The putative *CFO* orthologue *EpMADS17* was expressed in the peduncle, flower, fruit and protocorm (Figure 5). In rice, *CFO1* (*OsMADS32*) is only expressed in the panicle (Sang *et al.*, 2012). Orthologues of this gene have not been identified in other monocots. The limited range of expression of this gene may have prevented its identification in other orchids.

In the *SVP* subfamily, *EpMADS18* was expressed in all tissues except the protocorm and was highly expressed in fruit (Figure 5). *EpMADS19* was expressed highly in vegetative organs and in the peduncle (Figure 5). In *Arabidopsis*, *AGL24* is expressed in all tissues and *SVP* is only expressed in vegetative tissue (Parenicová *et al.*, 2003). In rice, *OsMADS47* is weakly expressed in vegetative organs, *OsMADS55* is highly expressed in the leaf, and *OsMADS22* is only expressed in the root (Arora *et al.*, 2007). The five grape *VvSVP* genes are all highly expressed in the bud, with *VvSVP5* also highly expressed in the leaf and *VvSVP2* in the shoot (Díaz-Riquelme *et al.*, 2009). None of the four cucumber *SVP* genes are expressed in the flower (Hu and Liu, 2012), similar to the pattern observed for *EpMADS19,* which was expressed in the flower at a very low level.

In the *AG* subfamily, *EpMADS20/21/22* were highly expressed in fruits (Figure 5). *EpMADS23* is a D-class MADS-box gene belonging to the STK-like lineage and was also highly expressed in fruits and absent in vegetative tissue (Figure 5). Similar spatial expression patterns in the column, ovary and developing ovules have been observed in other orchids, including *Phalaenopsis* 'Hatsuyuki' (*PhalAG1*, *PhalAG2*), *Dendrobium thyrsiflorum* (*DthyrAG1*, *DthyrAG2*) and *Phalaenopsis* 'Athens' (*PhaMADS8*, *PhaMADS9*, *PhaMADS10*) (Acri-Nunes-Miranda and Mondragon-Palomino, 2014; Skipper *et al.*, 2006; Song *et al.*, 2006). Although these genes belong to different species, *PhalAG1, DthyrAG1, PhaMADS8* and *PhaMADS10* have been classified in the C-class of *AG*-like genes; *PhalAG2*, *DthyrAG2* and *PhaMADS9* have been classified in the D-class of *AG*-like genes (Acri-Nunes-Miranda and Mondragon-Palomino, 2014; Skipper *et al.*, 2006; Song *et al.*, 2006). Our data also suggest that the C- and D-class genes in *E. pusilla* act redundantly in floral and ovule development.

The *E. pusilla* B sister gene, *EpMADS24*, was highly expressed in fruits (Figure 5). In *Arabidopsis*, the B sister genes are also expressed in the inflorescence and root (Parenicová *et al.*, 2003). In rice, expression of *OsMADS29* and *OsMADS31* is observed only in seed, whereas *OsMADS30* is expressed in both the panicle and seed. In grape, *VvBS1* is highly expressed in the bud, whereas *VvBS2* is highly expressed in the flower (Díaz-Riquelme *et al.*, 2009).

Only one MIKC*-type MADS-box gene, *EpMADS25,* was identified in *E. pusilla* (Figure 5). *Arabidopsis* encodes 7 MIKC*-type genes with diverse expression patterns, none of which are expressed exclusively in the silique (Parenicová *et al.*, 2003). The expression patterns of rice MIKC*-type MADS-box genes are also diverse (Arora *et al.*, 2007). More data are needed to further support the role of *EpMADS25* in gametophyte development.

Two truncated genes, *EpMADS26* and *EpMADS27*, exhibited distinct expression patterns in our study. *EpMADS26* was strongly detected in fruit, whereas *EpMADS27* was highly expressed in the protocorm. *EpMADS28*, an Mγ gene, is the only type I gene identified thus far in *E. pusilla* and was highly expressed in fruits (Figure 5). Only a small fraction of the type I genes from *Arabidopsis* have been functionally characterized (Kohler *et al.*, 2003; Parenicová *et al.*, 2003; Portereiko *et al.*, 2006). More data are needed to further support the role of *EpMADS28* in orchid function.

Whereas many of the *MADS*-like genes identified in *E. pusilla* exhibit expression patterns similar to those in *Arabidopsis* or other plant species, there are also a number of genes with unique expression profiles. These data will aid the determination of the functions of these *EpMADSs* in the development of the orchid flower and floral organ identity. Both similarities and differences will be of interest because orchid flowers are more diverse than the simple *Arabidopsis* flower and differ from other monocot flowers such as rice.

Spatial expression of *EpMADS* in floral organs

To further characterize the function of *EpMADSs* in floral development, we examined the spatial expression of *EpMADSs* in the sepal (outer tepal), lateral petal and lip (inner tepal), column (contains stamen) and ovule by qRT-PCR (Figure 6). The *EpMADS* expression results were compared to those of known genes in the ABCDE model.

Among the *AP1-FUL* subfamily genes, *EpMADS10* and *EpMADS11* were expressed in all selected floral organs. All three gene transcripts were relatively strong in the ovary (Figure 6a). To further explore the roles of *EpMADS12* in floral development, we investigated the spatiotemporal expression patterns of *EpMADS12* in floral bud and fruits via *in situ* hybridization. *EpMADS12* was detected in the floral primordium, lip primordium, column, pollinarium and seed (Figure 6b,c). The observed expression patterns of *EpMADS10/11/12* were similar to those of *DOMADS1* transcripts in *Dendrobium* 'Madame Thong-In', which are detected in all floral organs and the inflorescence meristem and floral primordium (Aceto and Gaudio, 2011; Yu and Goh, 2000). *DOMADS2* transcripts can be detected early in the apical meristem during the floral transition in *Dendrobium* 'Madame Thong-In'; however, its expression is restricted to the column in later development, similar to that of *EpMADS12* in the same organ (Aceto and Gaudio, 2011; Yu and Goh, 2000).

Among the *AP3-PI* subfamily genes, *EpMADS13* was highly expressed in the lateral petal, lip and column but was not detected in the sepal or ovules (Figure 6a). This expression pattern is similar to that of the *AP3*-like paralog *PeMADS3*, although *PeMADS3* is strongly expressed in the inner tepal and lip and to a lesser extent in the column in *P. equestris* (Tsai *et al.*, 2005). *EpMADS14* mRNA was strongly detected in the outer and inner tepal (sepal and lateral petal) (Figure 6a). *EpMADS15* mRNA was strongly detected in the sepal and lateral petal and relatively weakly in the column and ovules (Figure 6a), similar to its paralogs *PeMADS2* and *OMADS5* (Chang *et al.*, 2010; Tsai *et al.*, 2005). *EpMADS15* was also expressed in the floral primordium, lip primordium, column, pollinarium, stigma and seed (Figure 6b, c). The expression pattern and level of *EpMADS16* were similar to

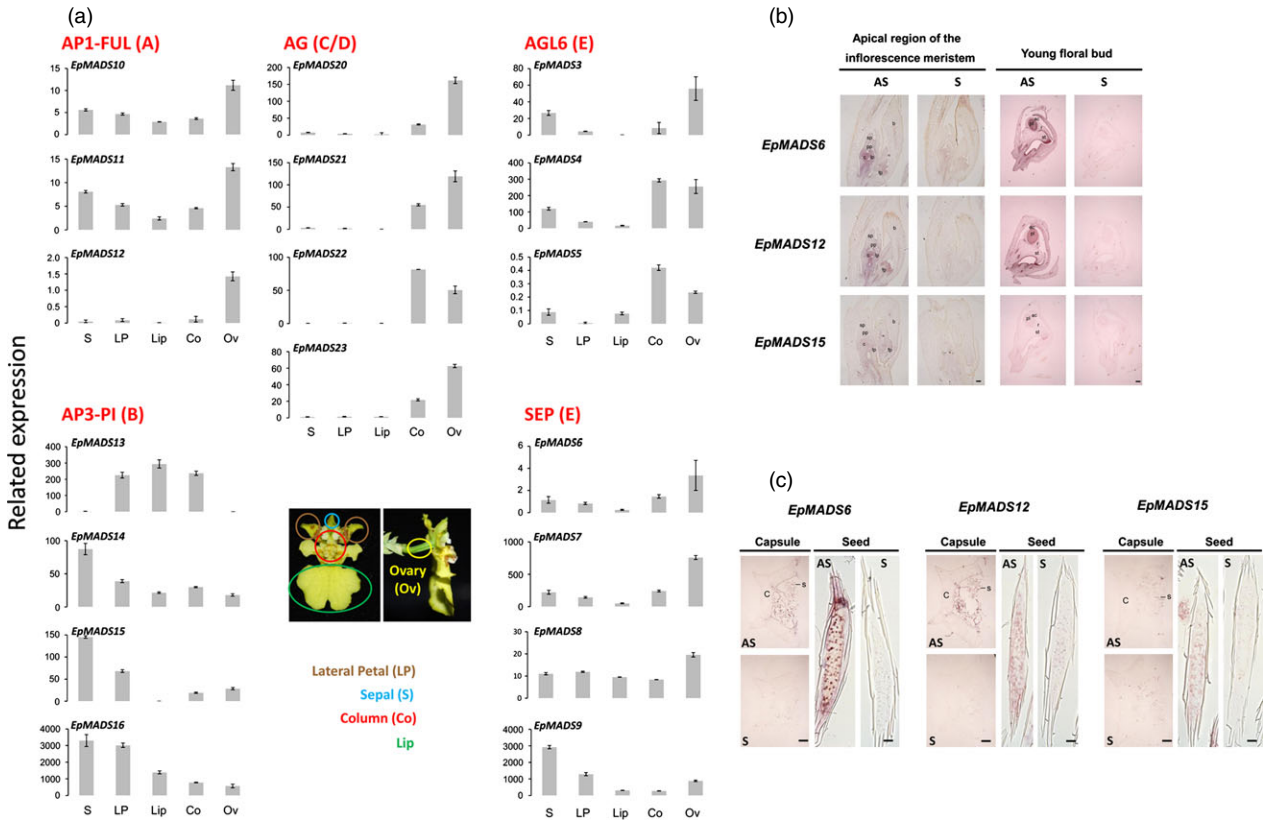

Figure 6 Floral organ-specific expression patterns of *EpMADS* genes. (a) Expression patterns of *EpMADS* genes from the specific organs shown in the photograph. RNA was extracted from five different tissues of *E. pusilla* and used for cDNA synthesis. S, sepal; Lp, lateral petal; Co, column; Ov, ovary. Real-time PCR was performed using the primers in Table S1. (b) *In situ* localization of *EpMADS1, 6, 12* and *15* transcripts in longitudinal sections in the inflorescence meristem (scale bar = 100 μm) and developing floral buds (scale bar = 200 μm). Sections were hybridized with an antisense 3′-specific RNA probe (antisense probe, AS) or the sense 3′-specific RNA probe (sense probe, S). Probes against the sense RNA fragment of *EpMADS1, 6, 12* and *15* serve as a negative control. ac, anther cap; b, bract; c, column; fp, floral primordium; lp, lip primordium; pp, petal primordium; pl, pollinarium; r, rostellum; sp, sepal primordium; st, stigma. (c) *In situ* localization of *EpMADS1, 6, 12* and *15* transcripts in longitudinal sections of fruits at 50 days after pollination (scale bar = 500 μm) and seed (scale bar = 10 μm). Sections were hybridized with an antisense 3′-specific RNA probe (antisense probe) or a sense 3′-specific RNA probe (sense probe). Probes against the sense RNA fragments of *EpMADS1, 6, 12* and *15* were used as negative controls. s, seed; c, capsule.

those of *EpMADS14*, albeit relatively weaker in the column and ovules of the flower (Figure 6a). *EpMADS16* belongs to the *PI/GLO*-like gene family and is similar to *PeMADS6* and *OMADS8*, which are expressed in the outer and inner tepal, lip, column and ovary (Chang *et al.*, 2010; Hsu and Yang, 2002; Tsai *et al.*, 2005). This expression pattern suggests that *EpMADS16* may be similar to *PeMADS6* and function in the development of these floral organs. The differential expression of paralogous *AP3* and *PI*-like MADS-box genes in Orchidaceae flowers and their sequence conservation in *E. pusilla* support the revised orchid code model (Mondragon-Palomino and Theißen, 2008, 2009, 2011).

In the *AG* subfamily genes, *EpMADS20/21/22* were strongly expressed in the column and ovules but were weakly detected in the sepal, lateral petal and lip (Figure 6a). In *Oncidium* 'Gower Ramsey', *OMADS4* is specifically expressed only in stamens and carpels (Hsu *et al.*, 2010). *PhalAG1* and *DcOAG1* are detected in all floral organs at the early stage of floral development; however, *PhalAG1* is also expressed in the column and lip at a later stage of floral development (Song *et al.*, 2006; Xu *et al.*, 2006). In *Phalaenopsis* 'Athens', the *AG*-like genes (*PhaMADS8*, *PhaMADS10*) are only expressed in the gynostemium and ovary (Acri-Nunes-Miranda and Mondragon-Palomino, 2014). These results suggest that *EpMADS20/21/22* are generally involved in stamen and carpel development, in accordance with other C-class genes such as *OMADS4*. The low level of expression of *EpMADS20/21* in the sepal indicates that a role for both genes in sepal development similar to that of *PhalAG1* in lip formation cannot be eliminated (Song *et al.*, 2006). More evidence is needed to define a functional role of *EpMADS20/21* in sepal development. High expression of *EpMADS23* was detected in the column and ovules (Figure 6a), similar to the C-class genes *EpMADS20/21/22* and *OMADS2* (Hsu *et al.*, 2010). These results suggest that *EpMADS23* is similar to other D-class orthologues in other orchid species, which are generally expressed in ovules (Figure 6a).

In the *SEP* subfamily, *EpMADS6/7/8/9* were expressed in all analysed organs (Figure 6a). *EpMADS6* was expressed in the floral primordium, lip primordium, column, pollinarium, stigma and seed, as detected by *in situ* hybridization (Figure 6b,c). In *Oncidium* 'Gower Ramsey', *OMADS6* and *OMADS11*, which are closely related to the *SEP3* and *SEP1/2* orthologues, respectively, are expressed in all four floral whorls, although in the stamens, *OMADS6* exhibits relatively low levels of expression, and

OMADS11 is absent (Chang *et al.*, 2009). In the orchid *Phalaenopsis*, *PhaMADS7* is differentially expressed in the inner lateral tepal and labellum (Acri-Nunes-Miranda and Mondragon-Palomino, 2014). Four *PeSEPs* are expressed in all floral organs (Pan *et al.*, 2014). These data suggest that *EpMADS6/7/8* are similar to most genes of the *AP1 FUL*-like subfamily, which are generally expressed in the early floral meristem and later floral organs during development but absent from vegetative tissues. However, some genes in monocots are also expressed in the leaf (Chen *et al.*, 2008; Pan *et al.*, 2014; Tzeng *et al.*, 2003). The other E-class genes in the AGL6 subfamily, *EpMADS3/4/5,* exhibited similar mRNA expression patterns and were strongly detected in the column and ovule (Figure 6a). Thus, the expression patterns of *EpMADS3/4/5* are similar to those of *OMADS7*, *AGL6* of *Arabidopsis* and *ZAG3* of *Z. mays*, with expression in all four flower organs and ovules (Chang *et al.*, 2009; Ma *et al.*, 1991; Mena *et al.*, 1995). *OMADS1* is expressed in the mature flower but is restricted to the lip and carpel, a pattern not observed for *EpMADS3/4/5*.

In contrast to *Arabidopsis* and other eudicots, the floral structure of orchids develops with extreme phenotypic differences between whorl 1 (outer tepal) and whorl 2 (inner tepal). Specifically, the lip in whorl 2 exhibits highly diversified morphology and colour among orchids, a feature that cannot be easily explained by the modified ABCDE model (Kanno *et al.*, 2003; Nakamura *et al.*, 2005). Based on the HOT and modified 'orchid code' models (Mondragon-Palomino and Theißen, 2011; Pan *et al.*, 2011) and the exploration of the expression patterns of *EpMADS* at the mRNA level, our findings also support the revised 'orchid code' hypothesis, which proposes that floral organ identity is regulated through the combined activity of duplicate MADS-box genes (Acri-Nunes-Miranda and Mondragon-Palomino, 2014; Mondragon-Palomino and Theißen, 2011). However, more data are required to define the precise mechanism by which these *EpMADS* genes control floral development in *E. pusilla*, including data from experiments such as *in situ* hybridization of *EpMADS* genes in floral organs, transcriptional activation of downstream targets and protein–protein interaction assays.

Subcellular localization of EpMADSs

The subcellular localization of EpMADS proteins was investigated via transient transformation of orchid flowers with yellow fluorescent protein (YFP) fused with one of the 28 EpMADS proteins. With the exception of proteins in the SOC and CFO families, the majority of the YFP-EpMADS signals were localized to nuclei (Figure 7). This result is similar to those for other plant MADS-box proteins and is consistent with the role of MADS-box proteins as transcription factors. Several MADS-box-containing transcription factors have been demonstrated to localize to the nucleus (Kaufmann *et al.*, 2005b; Shih *et al.*, 2014; Urbanus *et al.*, 2009; Verelst *et al.*, 2007; Zobell *et al.*, 2010). The motif KR[K/R]X4KK at positions 22–30 of the MADS-box domain plays an important role in the translocation of MADS-box proteins into the nucleus (Gramzow and Theißen, 2010). This NLS is present in all EpMADS proteins except the truncated EpMADS27.

SOC subfamily members EpMADS1 and EpMADS2 were localized throughout the cytosol (Figure 7). This result is similar to that for the *Arabidopsis* SOC1 protein (Lee *et al.*, 2008). The *Arabidopsis* genome contains six *SOC1/TM3-like* genes, including *SOC1* and two *STMADS11-like* genes (one of which is *AGL24*). During cotransformation with AGL24 and SOC1, SOC1 is translocated into the nucleus (Lee *et al.*, 2008). In bamboo (*Bambusa edulis*), when cotransformed with BeMADS1, BeMADS1 assists the translocation of BeMADS15, 16, 13, 21, 6 and 7 into the nuclei of lemma cells (Shih *et al.*, 2014). These results suggest that the subcellular localization of a MADS-box protein can be affected by protein–protein interaction with another MADS-box protein (Lee *et al.*, 2008; Shih *et al.*, 2014).

We used current *E. pusilla* research resources, including the transcriptome database and BAC library, to identify 28 MADS-box genes in this study. This is the first report to investigate the MADS-box gene family in this model orchid using pooled RNA sequencing via screening of NGS and BAC libraries. These fundamental resources in *E. pusilla* can be applied to commercial orchids. Because MADS-box proteins function as major determinants of floral organ identity, many of the *EpMADS* genes are expressed in reproductive tissues in patterns similar to those of other flowering-related MADS-box genes. The role of these *EpMADS* genes in floral organ development and flowering in *E. pusilla* is of great interest to both researchers and breeders working with commercial orchids. Our study is also the first to investigate the 'family-wide' subcellular localization of MADS-box genes in an orchid species. Because orchids are not closely related to other common model plants such as rice or *Arabidopsis*, it is paramount to develop specific resources for the study of orchids. These data specific for orchids will further both the basic biological understanding of Orchidaceae and selective commercial breeding of orchids.

Experimental procedures

RNA sample preparation

Erycina pusilla plantlets were incubated *in vitro* as described in Chiu *et al.* (2011). The root, leaf, peduncle, flower and fruit were separately harvested for total RNA extraction using TRIzol reagent (Invitrogen, Carlsbad, CA).

Identification of *EpMADS* genes

The *Arabidopsis* MADS-box transcription factor amino acid and nucleotide sequences were used to search for homologous unigenes in the *E. pusilla* unigene database using tBLASTn (Chou *et al.*, 2013; the unigene sequences are included in the supplemental file in this article). The 44 identified sequences were confirmed as MADS-box genes by BLASTx analysis. To obtain full-length genomic fragments, specific primers were designed to identify BAC clones containing the *EpMADS* genes (Table S1). The PCR strategy was described in Hsu *et al.* (2011) and Pan *et al.* (2012). The identified BAC plasmids were isolated for next-generation sequencing using the NucleoBond BAC 100 kit (Macherey-Nagel, Dueren, Germany). Five micrograms of plasmid was used for Illumina sequencing following the method described by Pan *et al.* (2012). The cDNA and genomic DNA sequences of the 28 *EpMADS* genes have been deposited in GenBank under the accession numbers KJ002726 to KJ002752, KJ766986 and KJ715209 to KJ715236.

Phylogenetic analysis

The deduced amino acid sequences of the *Arabidopsis* and orchid MADS-box genes were retrieved using the NCBI server (Table S2; http://www.ncbi.nlm.nih.gov/). Sequences in FASTA format were aligned using CLUSTAL W (Thompson *et al.*, 1994). Multiple sequence alignment and phylogenetic and molecular evolutionary

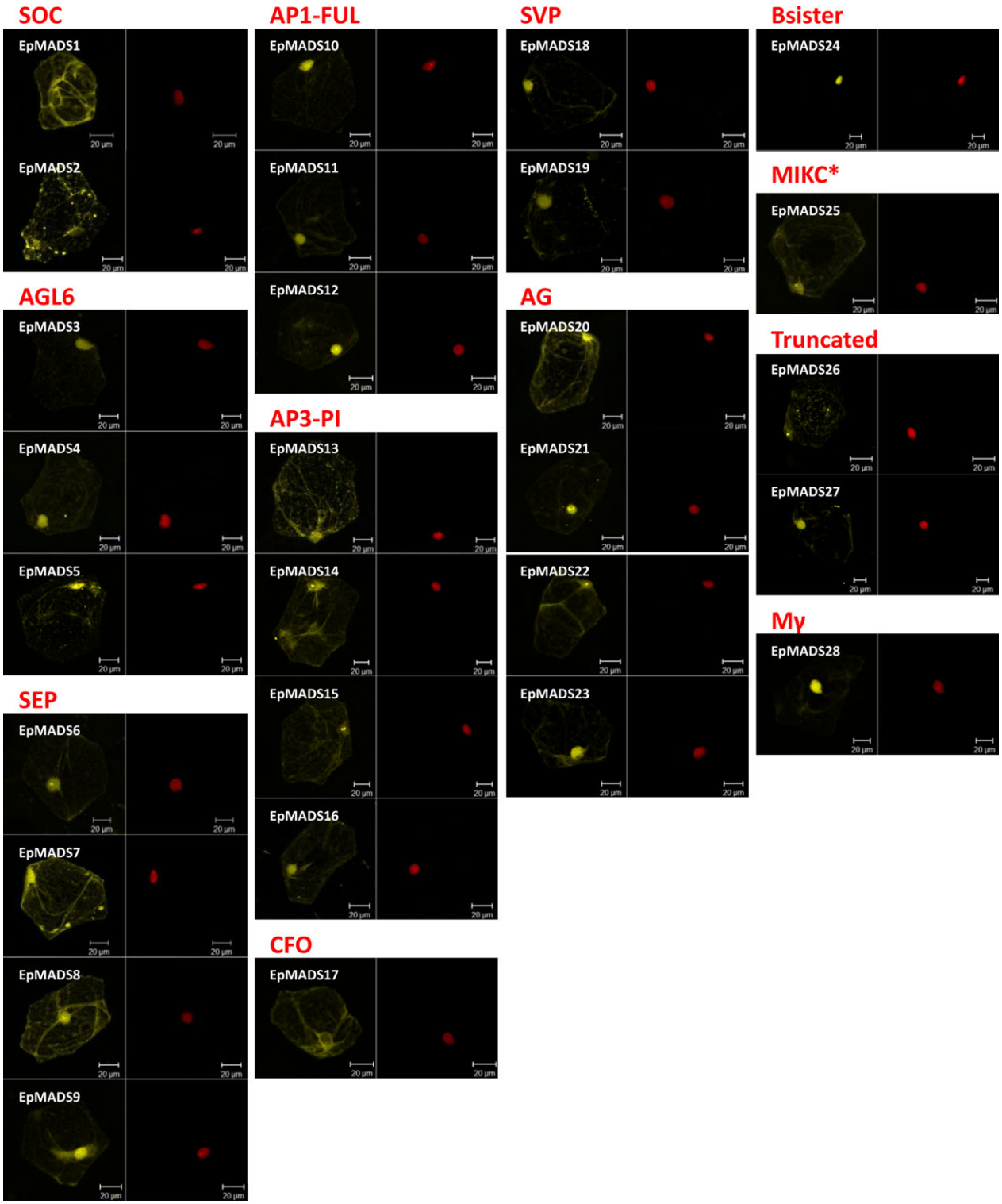

Figure 7 Subcellular localization of EpMADS fused with yellow fluorescent protein. Plasmids containing fusions of YFP and EpMADS driven by the CaMV35S promoter were transiently expressed in orchid petals according to Hsu *et al.* (2011) via delivery by particle bombardment (upper left panels). The NLS domain of VirD2 fused with mCherry was used to mark nuclei. Scale bar = 20 µm.

analyses were performed using MEGA software version 4 (Tamura *et al.*, 2007). The evolutionary history was inferred using the neighbour-joining (NJ) method after alignment (Altschul *et al.*, 1990). The distance matrices for the aligned sequences with all gaps ignored were calculated using the Kimura 2-parameter method (Kimura, 1980). The bootstrap consensus tree inferred from 1000 replicates was used to represent the evolutionary history of the genes analysed. All positions containing gaps and missing data were eliminated from the data set using the complete deletion option.

Domain identification

The MEME server was used to predict potential protein motifs (Bailey *et al.*, 2009). The motifs discovered by MEME with expected values lower than 2^{e-30} were searched against the InterPro database with InterProScan (Mulder *et al.*, 2005) for further confirmation or elimination.

Real-time quantitative reverse transcription (qRT)-PCR

Total RNA (5 μg) extracted from various tissues, including root, leaf, peduncle, fruits, flower, sepal, lateral petal, lip and column, was subjected to qRT-PCR. The expression level of a target gene was detected with SYBR Green (Applied Biosystems, Waltham, MA) real-time PCR on Rotor-Gene Q real-time thermocyclers (Corbett Research, Germantown, MD, Australia). Data were analysed using Rotor-Gene Q software version 2.0 (Corbett Research, Germantown, MD) and Microsoft Excel (Microsoft, Redmond, WA). The primer sets are given in Table S1. Ubiquitin was used as the internal control. The relative expression level of the target gene compared with that of Actin was defined as—$\Delta CT = -[CT_{Target} - CT_{Ubiquitin}]$. The target/ubiquitin mRNA ratio was calculated as $2\text{-}\Delta CT$. Experiments were performed three times in triplicate.

In situ hybridization

Erycina pusilla inflorescence meristem and developing floral buds were collected and fixed immediately at 4 °C overnight. *In situ* hybridization was performed as previously described (Lin *et al.*, 2014). Sense and antisense probes were synthesized using a SP6/T7 digoxigenin RNA labelling kit according to the manufacturer's instructions (Roche, Indianapolis, IN). Hybridization signals were detected using an NBT/BCIP detection kit (Roche, Indianapolis, IN). After *in situ* hybridization, images of tissue sections were acquired using a Zeiss Axio Scope A1 microscope equipped with an AxioCam HRc camera (Zeiss, Oberkochen, Germany).

Transient transformation and subcellular localization

Full-length cDNAs were amplified from *E. pusilla* cDNA via PCR. Products were cloned into vector pDONR221 using Gateway BP Clonase II Enzyme Mix (Invitrogen) and into vector p2GWF7 (nYFP) using Gateway LR Clonase II Enzyme Mix (Invitrogen, Karimi *et al.*, 2002). Plasmids (2.5 μg) were isolated and used to transform orchid floral lips by bombardment transformation (Hsu *et al.*, 2011). The transformed lips were incubated overnight and observed on a Zeiss LSM 510 META laser-scanning confocal microscope using an LD C-Apochromat 40 × /1.1 W objective lens (Hsu *et al.*, 2011).

Acknowledgements

We thank Ms. Anita K. Snyder for providing comments on the manuscript. We thank Sandy Chen for rice, *Arabidopsis* and orchid phylogenetic analysis. We also thank Hui-Ting Yang and Yi-Chen Lien for assistance with *in situ* hybridization. This work was supported by Academia Sinica and Innovative Translational Agricultural Research Administrative Office.

Author contributions

CSL, MLC, MTC and MCS conceived and designed the experiments. CTH, DCL and WJC performed the experiments. SSK performed *in situ* hybridization. JJWC, YTH, and MCS performed NGS and Bioinformatic analysis. JJW, YTH, MTC and MCS contributed reagents/materials/analysis tools. CSL, MLC and MCS wrote the manuscript.

Disclosure

The authors have no conflict of interests to declare.

References

Aceto, S. and Gaudio, L. (2011) The MADS and the beauty: genes involved in the development of orchid flowers. *Curr. Genomics*, **12**, 342–356.

Acri-Nunes-Miranda, R. and Mondragon-Palomino, M. (2014) Expression of paralogous *SEP*-, *FUL*-, *AG*- and *STK*-like MADS-box genes in wild-type and peloric *Phalaenopsis* flowers. *Front. Plant Sci.* **5**, 76.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Arora, R., Agarwal, P., Ray, S., Singh, A.K., Singh, V.P., Tyagi, A.K. and 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, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W. and Noble, W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208.

Becker, A., Winter, K.U., Meyer, B., Saedler, H. and Theißen, G. (2000) MADS-box gene diversity in seed plants 300 million years ago. *Mol. Biol. Evol.* **17**, 1425–1434.

Bemer, M., Gordon, J., Weterings, K. and Angenent, G.C. (2010) Divergence of recently duplicated Mγ-type MADS-box genes in Petunia. *Mol. Biol. Evol.* **27**, 481–495.

Chandler, S.F. and Sanchez, C. (2012) Genetic modification; the development of transgenic ornamental plant varieties. *Plant Biotechnol. J.* **10**, 891–903.

Chang, Y.Y., Chiu, Y.F., Wu, J.W. and Yang, C.H. (2009) Four orchid (*Oncidium* Gower Ramsey) *AP1/AGL9*-like MADS box genes show novel expression patterns and cause different effects on floral transition and formation in *Arabidopsis thaliana*. *Plant Cell Physiol.* **50**, 1425–1438.

Chang, Y.Y., Kao, N.H., Li, J.Y., Hsu, W.H., Liang, Y.L., Wu, J.W. and Yang, C.H. (2010) Characterization of the possible roles for B class MADS box genes in regulation of perianth formation in orchid. *Plant Physiol.* **152**, 837–853.

Chen, D., Guo, B., Hexige, S., Zhang, T., Shen, D. and Ming, F. (2007) *SQUA*-like genes in the orchid *Phalaenopsis* are expressed in both vegetative and reproductive tissues. *Planta*, **226**, 369–380.

Chen, M.K., Lin, I.C. and Yang, C.H. (2008) Functional analysis of three lily (*Lilium longiflorum*) *APETALA1*-like MADS box genes in regulating floral transition and formation. *Plant Cell Physiol.* **49**, 704–717.

Chen, Y.Y., Lee, P.F., Hsiao, Y.Y., Wu, W.L., Pan, Z.J., Lee, Y.I., Liu, K.W., Chen, L.J., Liu, Z.J. and Tsai, W.C. (2012) C- and D-class MADS-box genes from *Phalaenopsis equestris* (Orchidaceae) display functions in gynostemium and ovule development. *Plant Cell Physiol.* **53**, 1053–1067.

Chiu, Y.T., Lin, C.S. and Chang, C. (2011) *In vitro* fruiting and seed production in *Erycina pusilla* (L.) *N. H. Willams and M. W. Chase*. *Propag. Ornam. Plants*, **11**, 131–136.

Chou, M.L., Shih, M.C., Chan, M.T., Liao, S.Y., Hsu, C.T., Haung, Y.T., Chen, J.J., Liao, D.C., Wu, F.H. and Lin, C.S. (2013) Global transcriptome analysis and identification of a *CONSTANS*-like gene family in the orchid *Erycina pusilla*. *Planta*, **237**, 1425–1441.

Coen, E.S. and Meyerowitz, E.M. (1991) The war of the whorls: genetic interactions controlling flower development. *Nature*, **353**, 31–37.

Colombo, L., Franken, J., Koetje, E., van Went, J., Dons, H.J., Angenent, G.C. and van Tunen, A.J. (1995) The petunia MADS box gene *FBP11* determines ovule identity. *Plant Cell*, **7**, 1859–1868.

Díaz-Riquelme, J., Lijavetzky, D., Martínez-Zapater, J.M. and Carmona, M.J. (2009) Genome-wide analysis of $MIKC^{C}$-type MADS box genes in grapevine. *Plant Physiol.* **149**, 354–369.

Ding, L., Wang, Y. and Yu, H. (2013) Overexpression of *DOSOC1*, an ortholog of Arabidopsis *SOC1*, promotes flowering in the orchid *Dendrobium* Chao Parya Smile. *Plant Cell Physiol.* **54**, 595–608.

Erdmann, R., Gramzow, L., Melzer, R., Theißen, G. and Becker, A. (2010) *GORDITA* (*AGL63*) is a young paralog of the *Arabidopsis thaliana* B*sister* MADS box gene *ABS* (*TT16*) that has undergone neofunctionalization. *Plant J.* **63**, 914–924.

Gramzow, L. and Theißen, G. (2010) A hitchhiker's guide to the MADS world of plants. *Genome Biol.* **11**, 214.

Gramzow, L., Ritz, M.S. and Theißen, G. (2010) On the origin of MADS-domain transcription factors. *Trends Genet.* **26**, 149–153.

Hartmann, U., Höhmann, S., Nettesheim, K., Wisman, E., Saedler, H. and Huijser, P. (2000) Molecular cloning of *SVP*: a negative regulator of the floral transition in *Arabidopsis*. *Plant J.* **21**, 351–360.

Henschel, K., Kofuji, R., Hasebe, M., Saedler, H., Münster, T. and 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.

Hepworth, S.R., Valverde, F., Ravenscroft, D., Mouradov, A. and Coupland, G. (2002) Antagonistic regulation of flowering-time gene *SOC1* by CONSTANS and FLC via separate promoter motifs. *EMBO J.* **21**, 4327–4337.

Hsiao, Y.Y., Pan, Z.J., Hsu, C.C., Yang, Y.P., Hsu, Y.C., Chuang, Y.C., Shih, H.H., Chen, W.H., Tsai, W.C. and Chen, H.H. (2011) Research on orchid biology and biotechnology. *Plant Cell Physiol.* **52**, 1467–1486.

Hsu, H.F. and Yang, C.H. (2002) An orchid (*Oncidium* Gower Ramsey) *AP3*-like MADS gene regulates floral formation and initiation. *Plant Cell Physiol.* **43**, 1198–1209.

Hsu, H.F., Huang, C.H., Chou, L.T. and Yang, C.H. (2003) Ectopic expression of an orchid (*Oncidium* Gower Ramsey) *AGL6*-like gene promotes flowering by activating flowering time genes in *Arabidopsis thaliana*. *Plant Cell Physiol.* **44**, 783–794.

Hsu, H.F., Hsieh, W.P., Chen, M.K., Chang, Y.Y. and Yang, C.H. (2010) C/D class MADS box genes from two monocots, orchid (*Oncidium* Gower Ramsey) and lily (*Lilium longiflorum*), exhibit different effects on floral transition and formation in *Arabidopsis thaliana*. *Plant Cell Physiol.* **51**, 1029–1045.

Hsu, C.T., Liao, D.C., Wu, F.H., Liu, N.T., Shen, S.C., Chou, S.J., Tung, S.Y., Yang, C.H., Chan, M.T. and Lin, C.S. (2011) Integration of molecular biology tools for identifying promoters and genes abundantly expressed in flowers of *Oncidium* Gower Ramsey. *BMC Plant Biol.* **11**, 60–74.

Hu, L. and Liu, S. (2012) Genome-wide analysis of the MADS-box gene family in cucumber. *Genome*, **55**, 245–256.

Kanno, A., Saeki, H., Kameya, T., Saedler, H. and Theißen, G. (2003) Heterotopic expression of class B floral homeotic genes supports a modified ABC model for tulip (*Tulipa gesneriana*). *Plant Mol. Biol.* **52**, 831–841.

Karimi, M., Inzé, D. and Depicker, A. (2002) GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* **7**, 193–195.

Kaufmann, K., Anfang, N., Saedler, H. and Theissen, G. (2005a) Mutant analysis, protein–protein interactions and subcellular localization of the Arabidopsis B*sister* (ABS) protein. *Mol. Genet. Genomics*, **274**, 103–118.

Kaufmann, K., Melzer, R. and Theißen, G. (2005b) MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. *Gene*, **347**, 183–198.

Kim, S.Y., Yun, P.Y., Fukuda, T., Ochiai, T., Yokoyama, J., Kameya, T. and Kanno, A. (2007) Expression of a *DEFICIENS*-like gene correlates with the differentiation between sepal and petal in the orchid, *Habenaria radiate* (Orchidaceae). *Plant Sci.* **172**, 319–326.

Kimura, M. (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120.

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

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

Lee, J. and Lee, I. (2010) Regulation and function of SOC1, a flowering pathway integrator. *J. Exp. Bot.* **61**, 2247–2254.

Lee, J., Oh, M., Park, H. and Lee, I. (2008) SOC1 translocated to the nucleus by interaction with AGL24 directly regulates *LEAFY*. *Plant J.* **55**, 832–843.

Lee, S.H., Li, C.W., Liau, C.H., Chang, P.Y., Liao, L.J., Lin, C.S. and Chan, M.-T. (2015) Establishment of an agrobacterium-mediated genetic transformation procedure for the experimental model orchid *Erycina pusilla*. *Plant Cell Tissue Organ Cult.* **120**, 211–220.

Li, X., Luo, J., Yan, T., Xiang, L., Jin, F., Qin, D., Sun, C. and Xie, M. (2013) Deep sequencing-based analysis of the *Cymbidium ensifolium* floral transcriptome. *PLoS ONE*, **8**, e85480.

Lin, C.S., Hsu, C.T., Lu, H.C., Chou, M.L., Liao, D.C., Huang, Y.T., Yeh, Y.Y., Chang, S.B., Shen, S.C., Wu, F.H., Chen, J.J.W., Shih, M.C. and Chan, M.T. (2013) Catalog of *Erycina pusilla* miRNA and categorization of reproductive phase-related miRNAs and their target gene families. *Plant Mol. Biol.* **82**, 193–204.

Lin, H.Y., Chen, J.C., Wei, M.J., Lien, Y.C., Li, H.H., Ko, S.S., Liu, Z.H. and Fang, S.C. (2014) Genome-wide annotation, expression profiling, and protein interaction studies of the core cell-cycle genes in *Phalaenopsis aphrodite*. *Plant Mol. Biol.* **84**, 203–226.

Liu, Z. and Mara, C. (2010) Regulatory mechanisms for floral homeotic gene expression. *Semin. Cell Dev. Biol.* **21**, 80–86.

Ma, H. and dePamphilis, C. (2000) The ABCs of floral evolution. *Cell*, **101**, 5–8.

Ma, H., Yanofsky, M.F. and Meyerowitz, E.M. (1991) *AGL1-AGL6*, an *Arabidopsis* gene family with similarity to floral homeotic and transcription factor genes. *Genes Dev.* **5**, 484–495.

Malcomber, S.T. and Kellogg, E.A. (2005) *SEPALLATA* gene diversification: brave new whorls. *Trends Plant Sci.* **10**, 427–435.

Mena, M., Mandel, M.A., Lerner, D.R., Yanofsky, M.F. and Schmidt, R.J. (1995) A characterization of the MADS-box gene family in maize. *Plant J.* **8**, 845–854.

Michaels, S.D. and Amasino, R.M. (1999) *FLOWERING LOCUS C* encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*, **11**, 949–956.

Mondragon-Palomino, M. (2013) Perspectives on MADS-box expression during orchid flower evolution and development. *Front. Plant Sci.* **4**, 377.

Mondragon-Palomino, M. and Theißen, G. (2008) MADS about the evolution of orchid flowers. *Trends Plant Sci.* **13**, 51–59.

Mondragon-Palomino, M. and Theißen, G. (2009) Why are orchid flowers so diverse? Reduction of evolutionary constraints by paralogues of class B floral homeotic genes. *Ann. Bot.* **104**, 583.

Mondragon-Palomino, M. and Theißen, G. (2011) Conserved differential expression of paralogous *DEFICIENS*- and *GLOBOSA*-like MADS-box genes in the flowers of Orchidaceae: refining the 'orchid code'. *Plant J.* **66**, 1008–1019.

Mulder, N.J., Apweiler, R., Attwood, T.K., Bairoch, A., Bateman, A., Binns, D., Bradley, P., Bork, P., Bucher, P., Cerutti, L., Copley, R., Courcelle, E., Das, U., Durbin, R., Fleischmann, W., Gough, J., Haft, D., Harte, N., Hulo, N., Kahn, D., Kanapin, A., Krestyaninova, M., Lonsdale, D., Lopez, R., Letunic, I., Madera, M., Maslen, J., McDowall, J., Mitchell, A., Nikolskaya, A.N., Orchard, S., Pagni, M., Ponting, C.P., Quevillon, E., Selengut, J., Sigrist, C.J., Silventoinen, V., Studholme, D.J., Vaughan, R. and Wu, C.H. (2005) InterPro, progress and status in 2005. *Nucleic Acids Res.* **33**, D201–D205.

Nakamura, T., Fukuda, T., Nakano, M., Hasebe, M., Kameya, T. and Kanno, A. (2005) The modified ABC model explains the development of the petaloid perianth of *Agapanthus praecox* ssp. *orientalis* (Agapanthaceae) flowers. *Plant Mol. Biol.* **58**, 435–445.

Nam, J., Kim, J., Lee, S., An, G., Ma, H. and Nei, M. (2004) Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms. *Proc. Natl Acad. Sci. USA*, **101**, 1910–1915.

Ng, M. and Yanofsky, M. (2001) Function and evolution of the plant MADS-box gene family. *Nat. Rev. Genet.* **2**, 186–195.

Pan, Z.J., Cheng, C.C., Tsai, W.C., Chung, M.C., Chen, W.H., Hu, J.M. and Chen, H.H. (2011) The duplicated B-class MADS-box genes display dualistic characters in orchid floral organ identity and growth. *Plant Cell Physiol.* **52**, 1515–1531.

Pan, I.C., Liao, D.C., Wu, F.H., Daniell, H., Singh, N.D., Chang, C., Shih, M.C., Chan, M.T. and Lin, C.S. (2012) Complete chloroplast genome sequence of an orchid model plant candidate: *Erycina pusilla* apply in tropical *Oncidium* breeding. *PLoS ONE*, **7**, e34738.

Pan, Z.J., Chen, Y.Y., Du, J.S., Chen, Y.Y., Chung, M.C., Tsai, W.C., Wang, C.N. and Chen, H.H. (2014) Flower development of *Phalaenopsis* orchid

involves functionally divergent *SEPALLATA*-like genes. *New Phytol.* **202**, 1024–1042.

Parenicová, L., deFolter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., Angenent, G.C. and Colombo, L. (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.

Pinyopich, A., Ditta, G.S., Savidge, B., Liljegren, S.J., Baumann, E., Wisman, E. and Yanofsky, M.F. (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature*, **424**, 85–88.

Portereiko, M.F., Lloyd, A., Steffen, J.G., Punwani, J.A., Otsuga, D. and Drews, G.N. (2006) *AGL80* is required for central cell and endosperm development in *Arabidopsis*. *Plant Cell*, **18**, 1862–1872.

Prasad, K., Zhang, X., Tobón, E. and Ambrose, B.A. (2010) The Arabidopsis B-sister MADS-box protein, GORDITA, represses fruit growth and contributes to integument development. *Plant J.* **62**, 203–214.

Rouse, D.T., Sheldon, C.C., Bagnall, D.J., Peacock, W.J. and Dennis, E.S. (2002) FLC, a repressor of flowering, is regulated by genes in different inductive pathways. *Plant J.* **29**, 183–191.

Sage-Ono, K., Ozeki, Y., Hiyama, S., Higuchi, Y., Kamada, H., Mitsuda, N., Ohme-Takagi, M. and Ono, M. (2013) Induction of double flowers in *Pharbitis nil* using a class-C MADS box transcription factor with chimeric Repressor gene-silencing technology. *Plant Biotechnol.* **28**, 153–165.

Salemme, M., Sica, M., Gaudio, L. and Aceto, S. (2013) The *OitaAG* and *OitaSTK* genes of the orchid *Orchis italica*: a comparative analysis with other C- and D-class MADS-box genes. *Mol. Biol. Rep.* **40**, 3523–3535.

Sang, X., Li, Y., Luo, Z., Ren, D., Fang, L., Wang, N., Zhao, F., Ling, Y., Yang, Z., Liu, Y. and He, G. (2012) *CHIMERIC FLORAL ORGANS1*, encoding a monocot-specific MADS box protein, regulates floral organ identity in rice. *Plant Physiol.* **160**, 788–807.

Sedeek, K.E., Qi, W., Schauer, M.A., Gupta, A.K., Poveda, L., Xu, S., Liu, Z.J., Grossniklaus, U., Schiestl, F.P. and Schlüter, P.M. (2013) Transcriptome and proteome data reveal candidate genes for pollinator attraction in sexually deceptive orchids. *PLoS ONE*, **8**, e64621.

Shih, M.C., Chou, M.L., Yue, J.J., Hsu, C.T., Chang, W.J., Ko, S.S., Liao, D.C., Huang, Y.T., Chen, J.J., Yuan, J.L., Gu, X.P. and Lin, C.S. (2014) 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.* **14**, 179.

Skipper, M., Johansen, L.B., Pedersen, K.B., Frederiksen, S. and Johansen, B.B. (2006) Cloning and transcription analysis of an *AGAMOUS*- and *SEEDSTICK* ortholog in the orchid *Dendrobium thyrsiflorum* (Reichb. f.). *Gene*, **366**, 266–274.

Smaczniak, C., Immink, R.G., Angenent, G.C. and Kaufmann, K. (2012) Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. *Development*, **139**, 3081–3098.

Song, I.J., Nakamura, T., Fukuda, T., Yokoyama, J., Ito, T., Ichikawa, H., Horikawa, Y., Kameya, T. and Kanno, A. (2006) Spatiotemporal expression of duplicate *AGAMOUS* orthologues during floral development in *Phalaenopsis*. *Dev. Genes. Evol.* **216**, 301–313.

Su, C.L., Chao, Y.T., Alex Chang, Y.C., Chen, W.C., Chen, C.Y., Lee, A.Y., Hwa, K.T. and Shih, M.C. (2011) De novo assembly of expressed transcripts and global analysis of the *Phalaenopsis aphrodite* transcriptome. *Plant Cell Physiol.* **52**, 1501–1514.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* **24**, 1596–1599.

Tanaka, Y., Oshima, Y., Yamanura, T., Sugiyama, M., Mitsuda, N., Ohtsubo, N., Ohme-Takagi, M. and Terakawa, T. (2013) Multi-petal cyclamen flowers produced by AGAMOUS chimeric repressor expression. *Sci. Rep.* **3**, 2641.

Tapia-López, R., García-Ponce, B., Dubrovsky, J.G., Garay-Arroyo, A., Pérez-Ruíz, R.V., Kim, S.H., Acevedo, F., Pelaz, S. and Alvarez-Buylla, E.R. (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. and Saedler, H. (2001) Plant biology: floral quartets. *Nature*, **409**, 469–471.

Theißen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J.T., Münster, T., Winter, K.U. and Saedler, H. (2000) A short history of MADS-box genes in plants. *Plant Mol. Biol.* **42**, 115–149.

Thiruvengadam, M., Chung, I.M. and Yang, C.H. (2012) Overexpression of *Oncidium* MADS box (*OMADS1*) gene promotes early flowering in transgenic orchid (*Oncidium* Gower Ramsey). *Acta Physiol. Plant*, **34**, 1295–1302.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.

Tsai, W.C., Kuoh, C.S., Chuang, M.H., Chen, W.H. and Chen, H.H. (2004) Four *DEF*-like MADS box genes displayed distinct floral morphogenetic roles in *Phalaenopsis* orchid. *Plant Cell Physiol.* **46**, 831–844.

Tsai, W.C., Lee, P.F., Chen, H.I., Hsiao, Y.Y., Wei, W.J., Pan, Z.J., Chuang, M.H., Kuoh, C.S., Chen, W.H. and Chen, H.H. (2005) *PeMADS6*, a *GLOBOSA*/*PISTILLATA*-like gene in *Phalaenopsis equestris* involved in petaloid formation, and correlated with flower longevity and ovary development. *Plant Cell Physiol.* **46**, 1125–1139.

Tsai, W.C., Pan, Z.J., Hsiao, Y.Y., Jeng, M.F., Wu, T.F., Chen, W.H. and Chen, H.H. (2008) Interactions of B-class complex proteins involved in tepal development in *Phalaenopsis* orchid. *Plant Cell Physiol.* **49**, 814–824.

Tzeng, T.Y., Hsiao, C.C., Chi, P.J. and Yang, C.H. (2003) Two lily *SEPALLATA*-like genes cause different effects on floral formation and floral transition in Arabidopsis. *Plant Physiol.* **133**, 1091–1101.

Urbanus, S.L., de Folter, S., Shchennikova, A.V., Kaufmann, K., Immink, R.G. and Angenent, G.C. (2009) In planta localisation patterns of MADS domain proteins during floral development in *Arabidopsis thaliana*. *BMC Plant Biol.* **9**, 5.

Vaz, A.P.A., Figueiredo-Ribeiro, R.C.L. and Kerbauy, G.B. (2004) Photoperiod and temperature effects on in vitro growth and flowering of *P. pusilla* an epiphytic orchid. *Plant Physiol. Biochem.* **42**, 411–415.

Verelst, W., Saedler, H. and Münster, T. (2007) MIKC* MADS-protein complexes bind motifs enriched in the proximal region of late pollen-specific Arabidopsis promoters. *Plant Physiol.* **143**, 447–460.

Wang, J.W., Czech, B. and Weigel, D. (2009) miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell*, **138**, 738–749.

Weigel, D. and Meyerowitz, E.M. (1994) The ABCs of floral homeotic genes. *Cell*, **78**, 203–209.

Wu, R.M., Walton, E.F., Richardson, A.C., Wood, M., Hellens, R.P. and Varkonyi-Gasic, E. (2012) Conservation and divergence of four kiwifruit *SVP*-like MADS-box genes suggest distinct roles in kiwifruit bud dormancy and flowering. *J. Exp. Bot.* **63**, 797–807.

Xu, Y., Teo, L.L., Zhou, J., Kumar, P.P. and Yu, H. (2006) Floral organ identity genes in the orchid *Dendrobium crumenatum*. *Plant J.* **46**, 54–68.

Yu, H. and Goh, C.J. (2000) Identification and characterization of three orchid MADS-box genes of the AP1/AGL9 subfamily during floral transition. *Plant Physiol.* **123**, 1325–1336.

Yu, D., Kotilainen, M., Pöllänen, E., Mehto, M., Elomaa, P., Helariutta, Y., Albert, V.A. and Teeri, T.H. (1999) Organ identity genes and modified patterns of flower development in *Gerbera hybrida* (Asteraceae). *Plant J.* **17**, 51–62.

Yu, H., Xu, Y., Tan, E.L. and Kumar, P.P. (2002) *AGAMOUS-LIKE 24*, a dosage-dependent mediator of the flowering signals. *Proc. Natl Acad. Sci. USA*, **99**, 16336–16341.

Zahn, L.M., Kong, H., Leebens-Mack, J.H., Kim, S., Soltis, P.S., Landherr, L.L., Soltis, D.E., Depamphilis, C.W. and Ma, H. (2005) The evolution of the *SEPALLATA* subfamily of MADS-box genes: a preangiosperm origin with multiple duplications throughout angiosperm history. *Genetics*, **169**, 2209–2223.

Zhang, J., Wu, K., Zeng, S., Teixeira da Silva, J.A., Zhao, X., Tian, C.E., Xia, H. and Duan, J. (2013) Transcriptome analysis of *Cymbidium sinense* and its application to the identification of genes associated with floral development. *BMC Genom.* **14**, 279.

Zhao, Q., Weber, A.L., McMullen, M.D., Guill, K. and Doebley, J. (2011) MADS-box genes of maize: frequent targets of selection during domestication. *Genet. Res. (Camb)*, **93**, 65–75.

Zobell, O., Faigl, W., Saedler, H. and Münster, T. (2010) MIKC* MADS-box proteins: conserved regulators of the gametophytic generation of land plants. *Mol. Biol. Evol.* **27**, 1201–1211.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Phylogenetic analysis of rice, *Arabidopsis* and orchid MADS-box proteins.

Figure S2 Putative motifs and aligned sequences of EpMADS proteins.

Table S1 Primer sequences.

Table S2 Orchid MADS-box protein accession numbers and sequences used in phylogenetic analysis.

