Annals of Botany **114**: 1507–1515, 2014
doi:10.1093/aob/mcu124, available online at www.aob.oxfordjournals.org

ANNALS OF BOTANY
Founded 1887

PART OF A SPECIAL ISSUE ON FLOWER DEVELOPMENT

Analysis of the arabidopsis *REM* gene family predicts functions during flower development

Otho Mantegazza[1], Veronica Gregis[1], Marta Adelina Mendes[1], Piero Morandini[1,2], Márcio Alves-Ferreira[3], Camila M. Patreze[4], Sarah M. Nardeli[3], Martin M. Kater[1] and Lucia Colombo[1,2,*]

[1]*Dipartimento di BioScienze, Università degli Studi di Milano, Via Celoria 26, 20133 Milan, Italy,* [2]*Consiglio Nazionale delle Ricerche, Istituto di Biofisica, Via Celoria 26, 20133 Milan, Italy,* [3]*Universidade Federal do Rio de Janeiro, Instituto de Biologia, Dept de Genética, Av. Carlos Chagas Filho, 373, Cidade Universitária 21941-902–Rio de Janeiro, RJ, Brazil and* [4]*Universidade Federal do Estado do Rio de Janeiro, Instituto de Biociencias, Departamento de Botanica Av. Pasteur, no. 458, Urca, sala 306, 22290-255 Rio de Janeiro, RJ, Brazil*

** For correspondence. E-mail lucia.colombo@unimi.it*

Received: 19 February 2014 Returned for revision: 28 March 2014 Accepted: 29 April 2014 Published electronically: 6 July 2014

• **Background and Aims** The *REM* (*Reproductive Meristem*) gene family of *Arabidopsis thaliana* is part of the B3 DNA-binding domain superfamily. Despite the fact that several groups have worked on the *REM* genes for many years, little is known about the function of this transcription factor family. This study aims to identify a set of *REM* genes involved in flower development and to characterize their function.
• **Methods** In order to provide an overview of the *REM* gene family, a detailed expression analysis for all *REM* genes of *A. thaliana* was performed and combined with a meta-analysis of ChIP-sequencing and microarray experiments.
• **Key Results** Two sets of phylogenetically closely related *REM* genes, namely *REM23*, *REM24* and *REM25*, and *REM34*, *REM35* and *REM36*, were identified as possibly being involved in the early stages of flower development. Single- and double-mutant combinations were analysed for these genes, and no phenotypic effects were detected during flower development.
• **Conclusions** The data suggest that the *REM34*, *REM35* and *REM36* group is the most interesting one, as *REM34* is co-expressed with the floral meristem identity (FMI) genes, they are bound by AP1, SVP, AP3 and PI, and they are expressed in the floral meristem and during the earliest stages of flower development. However, it appears that high levels of functional redundancy may conceal the exact function of these transcription factor genes.

Key words: *Arabidopsis thaliana*, transcription factor, reproductive meristem, *REM* gene, flower development, B3 domain, floral meristem identity, FMI.

INTRODUCTION

Over the last 15 years, a large number of transcription factors expressed during flower development in *Arabidopsis thaliana* have been identified by genome-wide analysis; however, the function of most of them is still unknown. For example, in 2006, Wellmer *et al*. identified 222 transcription factors expressed differentially during early stages of flower development (Wellmer *et al*., 2006) and, in 2010 Kaufmann *et al*. identified 2300 genes bound by the key floral regulatory protein APETALA1 (AP1; Kaufmann *et al*., 2010). REM (Reproductive Meristem) transcription factor-encoding genes were recurrently identified in such genome-wide screens.

The *REM* gene family of *A. thaliana* is composed of 45 genes (Romanel *et al*., 2009). All REM factors contain the B3 DNA-binding domain, which is always present in at least one copy, often in two or three and sometimes up to seven copies. Moreover, *REM* genes are phylogenetically divergent and extensively duplicated, and are mostly located in clusters in the *Arabidopsis* genome (Fig. 1; Swaminathan *et al*., 2008; Romanel *et al*., 2009). The largest cluster is located on chromosome 4 containing nine *REM* genes (*REM34*, *REM35*, *REM36*, *REM37*, *REM38*, *REM39*, *REM40*, *REM41* and *REM42*) within 30 kb. This cluster is partially duplicated on chromosome 2 (*REM29*, *REM30*, *REM31* and *REM32*).

The B3 domain was first identified in the maize VIVIPAROUS1 (VP1) transcription factor (McCarty *et al*., 1991) and was shown to bind DNA cooperatively *in vitro* (Suzuki *et al*., 1997). The B3 domain transcription factors (including REMs) are specific to plants and widespread among angiosperms, gymnosperms, ferns, mosses, liverworts and green algae (Swaminathan *et al*., 2008). The transcription factor families that contain the B3 domain have been shown to play important roles in plant development and are LAV [LEC2 (LEAFY COTYLEDON 2)/ABI3 (ABSCISIC ACID INSENSITIVE 3)−VAL (VP1/ABI3-LIKE)], RAV (RELATED to ABI3/VP1), ARF (AUXIN RESPONSE FACTOR) and REM (Swaminathan *et al*., 2008; Wang *et al*., 2012). The REM family is the least studied and understood. The B3 domains of LAV, RAV and ARF bind DNA specifically; the reported recognition sequence is different among these three families and is TGTCTC for ARF proteins (Ulmasov *et al*., 1997), CATGCA for LAV (Ezcurra *et al*., 2000) and CACCTG for RAV (Kagaya *et al*., 1999). The B3 domains of REM proteins show a variability in the sequence and length of the loops proposed to be responsible for specificity in DNA binding (Romanel *et al*., 2009; King *et al*., 2013).

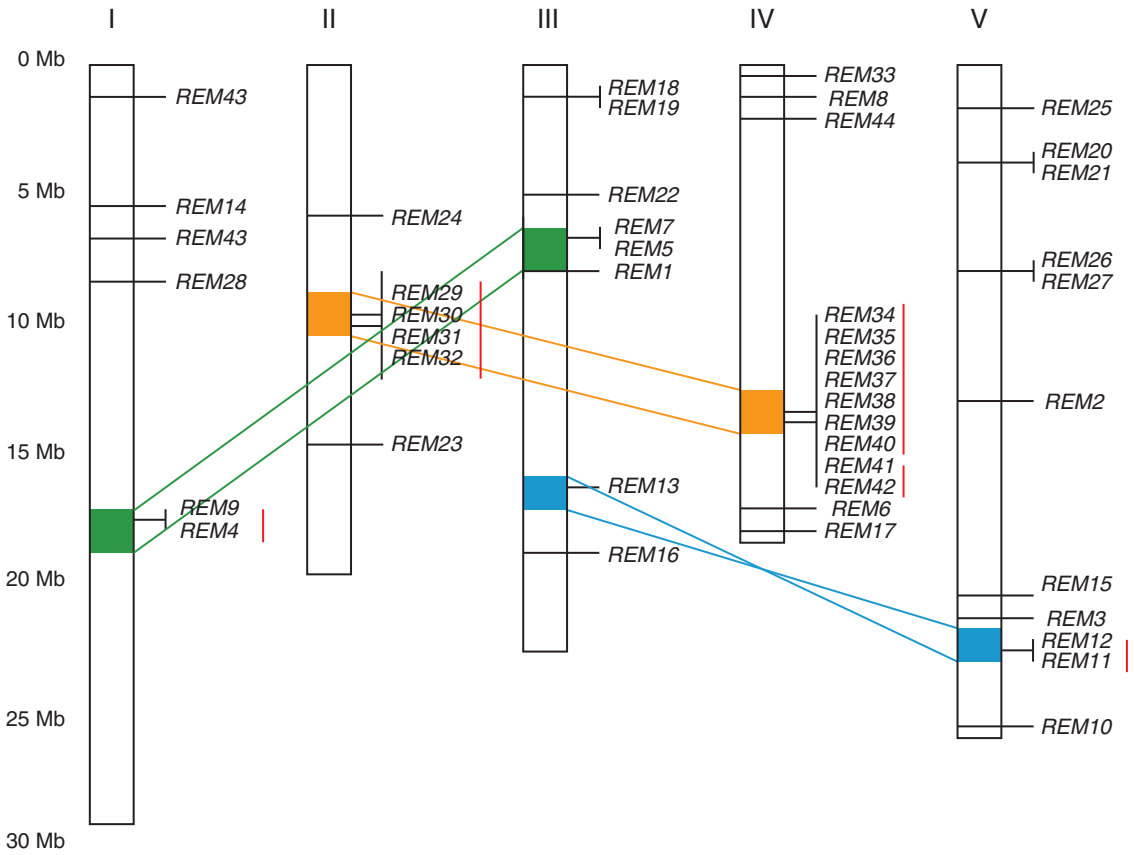

FIG. 1. Distribution of *REM* genes on *Arabidopsis* chromosomes indicating the phylogenetic relationships (coloured lines).

However, it is still unclear if REMs indeed bind DNA in a sequence-specific manner.

The first gene of this family that was identified in *Arabidopsis* is *REM34*, which was named *REM1* at that time. *REM34* was characterized as the putative orthologue of *BoREM1* of *Brassica oleracea* (Franco-Zorilla *et al*., 2002). BoREM1 was described as a novel regulatory protein putatively involved in the process of floral meristem identity determination. *BoREM1* is specifically expressed in the cauliflower curd of *B. oleracea* and was shown to have an expression pattern similar to that of the floral meristem identity (FMI) gene *BoLEAFY* (Franco-Zorrilla *et al*., 1999).

VERNALIZATION1/REM5 (*VRN1*) of *Arabidopsis* was the first *REM* gene to be functionally characterized. The *vrn1* mutant has a reduced vernalization response in long-day conditions (Levy *et al*., 2002). VRN1 promotes flowering and is proposed to be involved in the epigenetic repression of *FLOWERING LOCUS C* (Mylne *et al*., 2006). Moreover, overexpression of *VRN1* causes pleiotropic effects, and expression of a tagged version of this protein, acting as a dominant repressor, was lethal (Levy *et al*., 2002; King *et al*., 2013). It was hypothesized that *VRN1* is redundant with other genes involved in various processes since *VRN1* mRNA was detected in almost all plant tissues but the single mutant did not reveal phenotypes in all tissues (King *et al*., 2013). VRN1 is capable of DNA binding in a non-sequence-specific manner (Levy *et al*., 2002; Mylne *et al*., 2006; King *et al*., 2013).

VERDANDI/REM20 (VDD) is the only other REM transcription factor that is functionally characterized at present. In the *vdd* female gametophyte, the development of the antipodal and synergid cells is affected. *VDD* is a direct target of the MADS-domain ovule identity complex (Matias-Hernandez *et al*., 2010) that includes SEEDSTICK (STK), SEPALLATA3 (SEP3) and SHATTERPROOF1/2 (SHP1/2) (Favaro *et al*., 2003). STK and SEP3 regulate *VDD* expression during ovule development by binding to the *VDD* promoter at two different MADS-domain binding sites inducing a loop in the DNA (Mendes *et al*., 2013).

Several *REM* genes have distinctive expression patterns, but no information is available about their function. For example, *REM23* is expressed during early stages of flower development (Wellmer *et al*., 2006) while *REM22* has been used as a marker of early stamen development (Romanel *et al*., 2011).

Here we provide data about selected REM transcription factors that may play a role in flower development. Expression in different tissues and during flower development was investigated for all *REM* genes. Subsequently, in order to focus on those *REM* genes most likely to be involved in flower development, we performed a meta-analysis of available transcriptomic and ChIP-seq (chromatin immunoprecipitation coupled to next-generation sequencing) data. This resulted in the identification of two groups of *REM* genes (*REM23*, *REM24* and *REM25*, and *REM34*, *REM35* and *REM36*) compatible with a role in flower development. In order to define the function of *REM23*, *REM24* and

REM25 and *REM34*, *REM25* and *REM36*, single and double mutant combinations were analysed.

MATERIALS AND METHODS

Plant material and growth conditions

Plants of *Arabidopsis thaliana* were grown at 22 °C in short-day (8 h of light/16 h of dark) and long-day (16 h of light/8 h of dark) conditions. The insertion lines (FLAG_566H04, SAIL_753_B09, SAIL_116500, FLAG_089C09 and SALK_054142) were purchased from the NASC (Nottingham Arabidopsis Stock Center). To test if the insertion lines were complete knock-out, we tested for mRNA levels in inflorescence tissues by real-time quantitative PCR (qPCR). The primers used for genotyping and qPCR are listed in Supplementary Data Table S1.

RNA extraction and cDNA synthesis

We extracted RNA using the RNeasy Kit (Qiagen), following the manufacturer's instructions. After extraction, RNA samples (10 μg of RNA) were treated with RNase-free DNase I (Life Technologies), as specified by the manufacturer. RNA concentration and purity were determined using a NanoDropTM Spectrophotometer ND-1000 (Thermo Scientific). The integrity of the RNA was also assessed by 1 % agarose gel electrophoresis. The presence of spurious amplification products caused by genomic DNA was also continuously monitored by the verification of the qPCR dissociation profile. cDNAs were synthesized with Superscript™ III Reverse transcriptase (Life Technologies) as specified by the manufacturer.

Real-time quantitative PCR (qPCR)

Polymerase chain reactions were carried out in an optical 96-well plate with a 7500 Fast Real-Time PCR detection system (Applied Biosystems, Foster City, CA, USA), using SYBR®Green to monitor double-stranded DNA synthesis. Reaction mixtures contained 10 μL of diluted cDNA (1:50), 0·2 μM of each primer, 50 μM of each dNTP, 1 × PCR Buffer (Invitrogen), 3 mM $MgCl_2$, 2 μL of SYBR®Green I (Molecular Probes) diluted with water (1:10000), and 0·25 U of Platinum *Taq* DNA polymerase (Invitrogen), in a total volume of 20 μL. Reaction mixtures were incubated for 5 min at 94 °C, followed by 40 amplification cycles made up of: 15 s at 94 °C, 10 s at 60 °C and 15 s at 72 °C. The PCR efficiencies and the optimal quantification cycle threshold (C_T values) were estimated using the online real-time PCR Miner tool (Zhao and Fernald, 2005). For all genes studied, two independent biological samples of each experimental condition were evaluated using three technical replicates for each sample. The C_T value were converted into expression values with the ΔC_T method, a simplified version of the $\Delta\Delta C_T$ method as described by Livak and Schmittgen (2001); we used the gene At1g58050 as reference. Expression values were square root-transformed and visualized on a heatmap produced using MeV 3.1 (Multiple Experiment Viewer), which is freely available online.

Microarray-based co-expression analysis

Co-expression analysis was performed as described previously (Menges *et al*., 2007, 2008) using the expression values both untransformed and after transformation into logarithmic values. Heatmaps were produced using MeV 3.1. The AtGenExpress data (Schmid *et al*., 2005) were downloaded as 'absolute values' (linearized gcRMA values) from http://www.weigelworld.org/resources/microarray/AtGenExpress/. According to the website, the 'entire data set was quantile-normalized using gcRMA'. Each value is the mean of three replicates. The co-expression network was generated essentially as described by Berri *et al*. (2009).

In situ *hybridization*

In situ hybridization was performed as described previously (Gregis *et al*., 2009). Inflorescences were harvested in FAA [ethanol (Fluka) 50 %; acetic acid (Sigma-Aldrich) 5 %; formaldehyde (Sigma-Aldrich) 3·7 % (v/v)], infiltrated under mild vacuum conditions for 30 min, dehydrated in a graded ethanol series, transferred to bioclear (Bioptica) and then embedded in Paraplast X-tra® (Sigma-Aldrich). DNA fragments used to produce RNA probes were amplified from cDNA with the primers listed in Supplementary Data Table S1 and cloned in the pGEM®-T Easy vector (Promega). Digoxigenin-labelled antisense probes were transcribed and labelled from pGEM®-T Easy with T7 RNA polymerase (Promega) as specified by the manufacturer, and using the DIG RNA labelling mix (Roche). Paraplast-embedded tissues were sliced on an RM2155 microtome (Leica) and hybridized as described by Coen *et al*. (1990). Immunodetection was carried out with anti-digoxigenin-AP Fab fragment (Roche) and BCIP-NBT colour development substrate (Promega) as specified by the manufacturer.

RESULTS

Integrative analysis of expression patterns of REM *genes*

Based on the microarray data that are available in the NASC arrays database (ftp://arabidopsis.info/pub/NASCArrays/Data/), a heatmap of the expression patterns of *REM* genes was drawn (Fig. 2A). It should be noted that 11 *REM* genes (*REM9*, *REM30*, *REM31*, *REM32*, *REM36*, *REM38*, *REM39*, *REM40*, *REM42*, *REM43* and *REM45*) have no corresponding probe on the ATH1 array. Moreover, the ATH1 array probe 256918_s_at does not distinguish between *REM7* and *REM8*, and the same holds for probe 257436_s_at, which does not distinguish *REM29* and *REM33*; therefore, we refer to measurements from these probes as *REM7/8* and *REM29/33*. The microarray data reveal that only two *REM* genes are expressed in vegetative tissue, namely *REM4* and *REM5/VRN1*, whereas four *REM* genes are almost undetectable (*REM10*, *REM37*, *REM41* and *REM44*). The other *REM* genes are preferentially expressed during flower and seed development.

We decided to integrate and to refine the data provided by the microarray database using a qPCR approach, which is considered to be a 'gold standard' for expression analysis (Wang *et al*., 2006). The expression data obtained by these experiments are reported in a heatmap for easy interpretation (Fig. 2B). Notably, we also produced expression data for the genes that are not

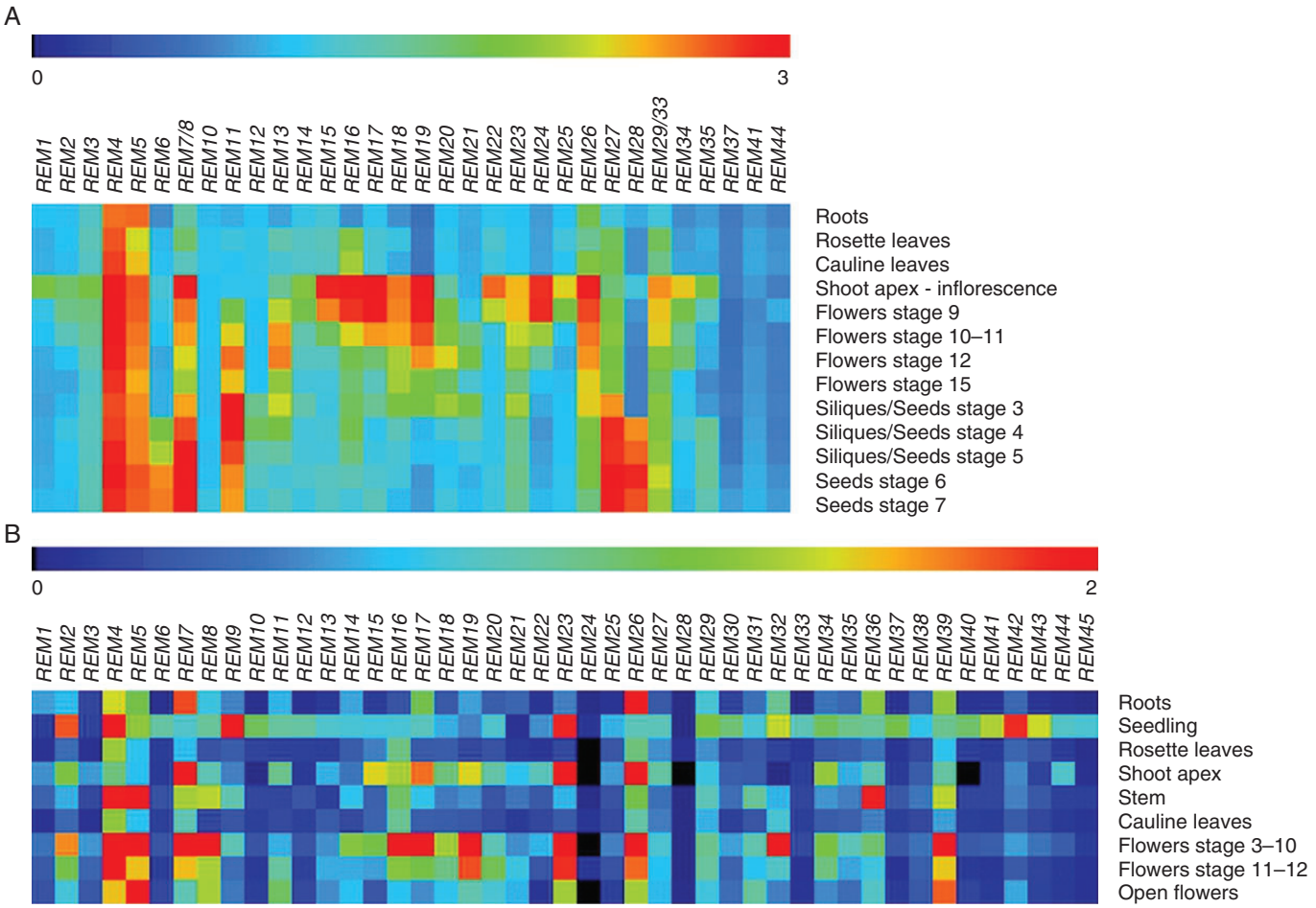

FIG. 2. Heatmaps for expression levels of *REM* genes. (A) *REM* gene expression levels measured by microarray; data are shown as $\log_{10}$ transformed absolute values from the AtGeneExpress developmental series; the 'Roots' sample corresponds to the ATGE_93 microarray experiment, the 'Rosette leaves' sample is the ATGE_13 experiment, the 'Cauline leaves' sample is the ATGE_26 experiment, the 'shoot apex - inflorescence' sample is the ATGE_29 experiment, the 'Flowers stage 9' sample is the ATGE_31 experiment, the 'Flowers stage 10–11' sample is the ATGE_32 experiment, the 'Flowers stage 12' sample is the ATGE_33 experiment, the 'Flowers stage 15' sample is the ATGE_39 experiment, the 'Silique/Seeds stage 3' sample is the ATGE_76 experiment, the 'Silique/Seeds stage 4' sample is the ATGE_77 experiment, the 'Silique/Seeds stage 5' sample is the ATGE_78 experiment, the 'Seeds stage 6' sample is the ATGE_79 experiment, and the 'Seeds stage 7' sample is the ATGE_81 experiment. (B) *REM* gene expression levels measured by quantitative real-time PCR (data are shown as square root-transformed $2^{-\Delta CT}$ values).

represented on the ATH1 microarray. Our analysis confirms the microarray data and defines expression patterns of *REM* genes over a wider quantitative range (Wang *et al.*, 2006) (note that expression data from qPCR are square root normalized for easier visualization on the heatmap). We confirmed that three groups of *REM* genes [(1) *REM4*, *REM5*, *REM7* and *REM8*; (2) *REM15*, *REM16*, *REM17*, *REM18* and *REM19*; and (3) *REM22* and *REM23*] are highly expressed during early stages of flower development, as suggested by the microarray data.

Co-expression of REM *genes with specific reproductive markers*

The analysis of microarray data to highlight co-expression of genes has yielded valuable predictions about gene functions in *Arabidopsis* as well as other organisms (Beekwilder *et al.*, 2008; Murgia *et al.*, 2011; for a review, see Usadel *et al.*, 2009). This approach is based on the 'guilt by association' concept, which suggests that genes with a similar expression 'behaviour' are likely to be involved in the same biological process.

Therefore, to mine in more depth and to test the involvement of specific *REM* genes in flower development, we analysed the degree of correlation between the expression profiles of *REM* genes and a selection of key floral regulators. We used an approach based on the calculation of the Pearson correlation coefficient (r) between the expression levels of arabidopsis genes measured in almost 2000 Affymetrix hybridizations (Menges *et al.*, 2007, 2008). The initial list of 45 *REM* genes was screened to remove those members that are missing from the Affymetrix array.

We analysed the correlation among *REM* genes and two groups of genes: (1) the FMI genes *LEAFY* (*LFY*), *AP1* and *CAULIFLOWER* (*CAL*) (Irish and Sussex, 1990; Weigel *et al.*, 1992); and (2) the MADS-box floral homeotic genes, *APETALA3* (*AP3*), *PISTILLATA* (*PI*), *AGAMOUS* (*AG*) and again *AP1* (for a review, see Krizek and Fletcher, 2005). As a positive control, we used *STK*, which directly regulates the expression of *REM20/VDD* (Matias-Hernandez *et al.*, 2010). Finally, as negative controls, we used *TERMINAL FLOWER 1* (*TFL1*), whose transcripts accumulate only in the inflorescence meristem and are excluded from the flower (Ratcliffe *et al.*, 1999), and

SHORT VEGETATIVE PHASE (*SVP*), which is an FMI gene but is also active in vegetative tissues (Gregis *et al*., 2008).

The correlation table is represented as a heatmap (Fig. 3; Supplementary Data Fig. S1) and as a graph (Fig. 4). The correlation values (*r*) of a larger gene set are reported in Supplementary Data Table S2. This analysis allows the definition of three large groups of *REM* genes, which become more evident after clustering. One large group of *REM* genes includes *REM16*, *REM17*, *REM18*, *REM19*, *REM22*, *REM23*, *REM24* and *REM34*, and showed good to very good correlation with the FMI gene *LFY* (most correlation coefficients are >0·7) and to a lesser degree with *AP1* and *CAL*. A second group that includes *REM1*, *REM4*, *REM7*, *REM8*, *REM26* and *REM29/33* showed a fair correlation with group 1, but they correlated more weakly with the FMI genes. A third group, comprising *REM11*, *REM13*, *REM20* and *REM21* (and to a lesser degree also *REM6*, *REM27* and

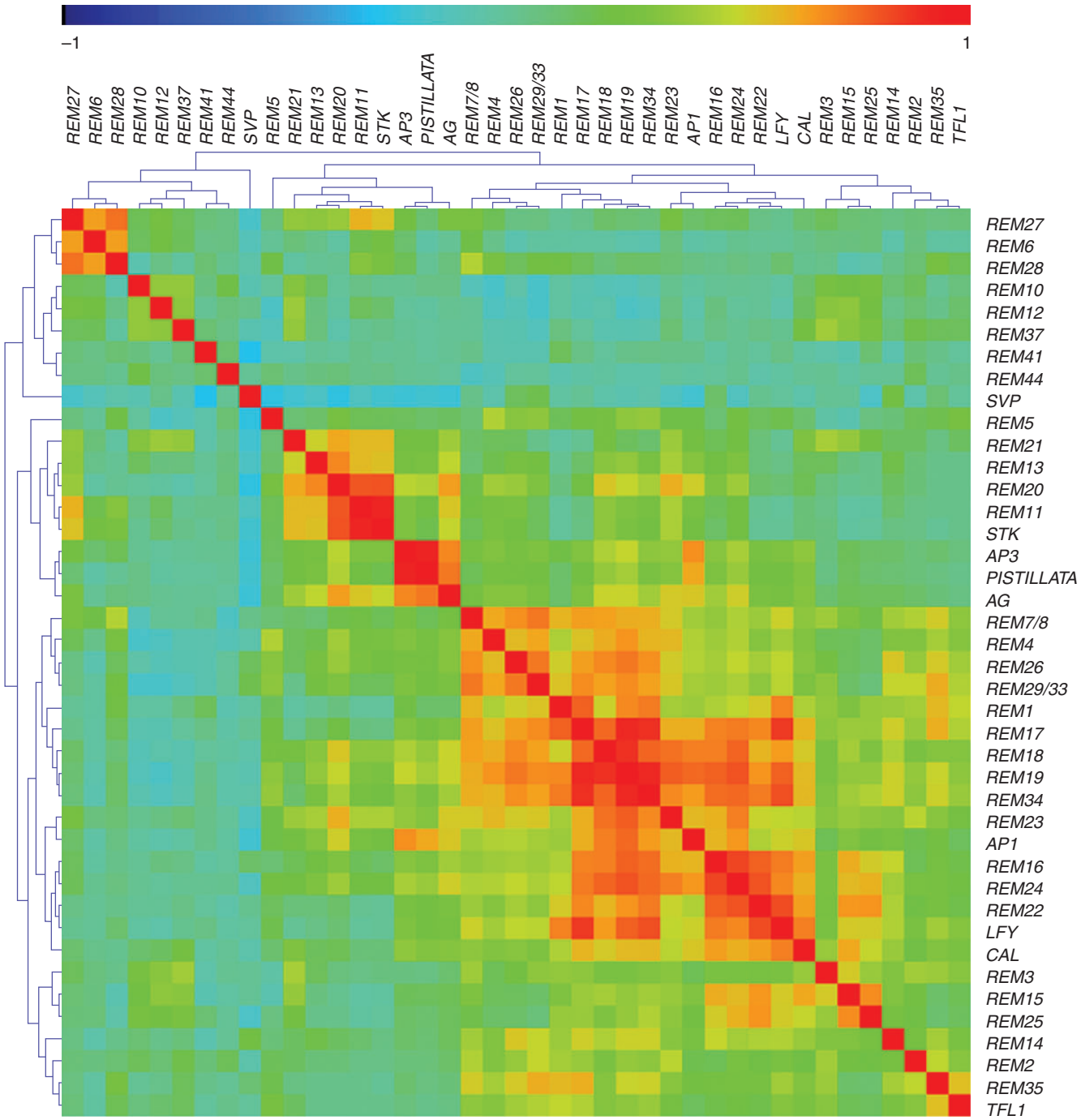

FIG. 3. Co-expression matrix for *REM* genes and key regulators of reproductive development (same data as in Fig 3) after hierarchical clustering with the Euclidean distance as metric and the average linkage method. The values visualized in the heatmap represent the Pearson correlation between the expression values of each gene pair calculated using almost 2000 microarray hybridization experiments.

REM28), showed good correlation with *STK*. We detected low correlation with the homeotic genes *AP3*, *PI* and *AG* ($0{\cdot}5 > r > -0{\cdot}5$), no significant correlation with *SVP* ($0{\cdot}2 > r > -0{\cdot}2$), and only one gene (*REM35*) correlates with *TFL1*, although weakly ($r = 0{\cdot}5627$). A high correlation between *STK* and *VDD* ($r = 0{\cdot}8116$) was evident (see Figs 3 and 4), confirming this positive control in our analysis.

REM *genes are the target of the floral identity transcription factors*

In order to investigate whether *REM* genes are under direct control of key regulators of flower development, we analysed available ChIP-seq data for binding of such factors to the putative regulatory regions of *REM* genes. In particular, we analysed the high confidence target data sets for LFY (Winter *et al*., 2011), SVP (Gregis *et al*., 2013), AP1 (Kaufmann *et al*., 2010), SEP3 (Kaufmann *et al*., 2009), PI and AP3 (Wuest *et al*., 2012), AG (Ó'Maoiléidigh *et al*., 2013), AGAMOUS-LIKE 15 (AGL15; Zheng *et al*., 2009) and AP2 (Yant *et al*., 2010). Interestingly, the ChIP-seq data are consistent with information obtained from the co-expression analysis. *REM17* and *REM18* are direct targets of LFY, and both genes were also co-expressed with *LFY*, and the same relationship was observed for AP1 and its direct target *REM34*. The complete list of direct targets can be found in Table 1. SEP3 and AP2 do not bind to any REM genomic region, whereas AP3 and PI were shown to have multiple binding sites in the *REM34*, *REM35* and *REM36* cluster on chromosome 4. These three genes are arranged on chromosome 4 within 10 kb in a tight cluster together with six other *REM* genes. Binding of AP3 and PI falls precisely in the short non-coding region between *REM35* and *REM36*, which is 565 bp from the stop codon of *REM36* to the start codon of *REM35*, and in the small intergenic region between *REM36* and *REM37*, which is 687 bp long from the stop codon of *REM37* to the start codon of *REM36* (see Fig. 5). *REM34*, *REM35* and *REM36* are upregulated in *pi-1* and *ap3–3* mutants (Wuest *et al*., 2012); in contrast, they do not change expression upon binding of AP1 (Kaufmann *et al*., 2010) and SVP (Gregis *et al*., 2013).

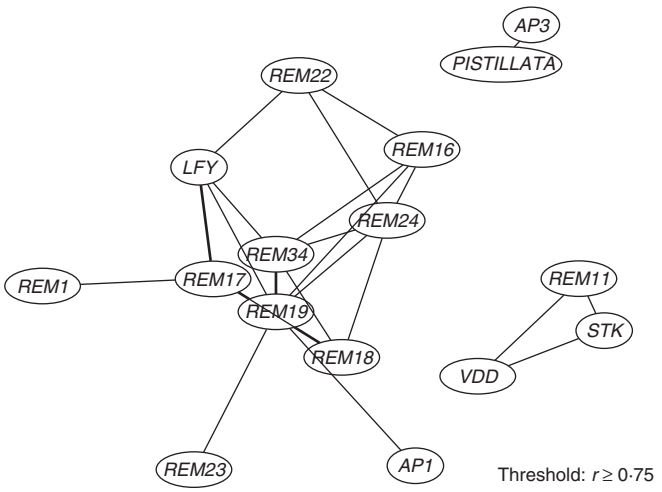

FIG. 4. Graphical representation of the co-expression among *REM* genes and selected reproductive markers. Only pairs of genes with co-expression levels (r) that are $\geq 0{\cdot}75$ are shown. Co-expressed genes are linked by lines, the length of the line being negatively correlated with co-expression levels.

Expression patterns of selected REM *genes during early stages of flower development*

We further characterized by RNA *in situ* hybridization experiments the expression profiles of *REM23*, *REM24*, *REM25*, *REM34*, *REM35* and *REM36* during early stages of flower development. We selected *REM23*, *REM24* and *REM25* because these genes are phylogenetically closely related (Romanel *et al*., 2009) and their expression shows correlation with the FMI genes. For the same reasons, we decided to characterize *REM34*, *REM35* and *REM36*, which are also direct targets of SVP, AP1, AP3 and PI (see above).

Our analysis showed that all the genes in the phylogenetic group of *REM24* had similar expression patterns (Fig. 6A–C), with *REM23*, *REM24* and *REM25* being expressed in stamens starting from stage 7/8 of flower development. None of them was detectable at earlier stages. The genes in the *REM34* phylogenetic group also showed similar expression patterns to each other (Fig. 6D–F), with *REM34*, *REM35* and *REM36* being

TABLE 1. *Complete set of* REM *genes that are direct targets of the key floral transcription factors AP1, AP3, PI, LEAFY and SVP*

TAIR ID	FEATURE ID	AP1 target	AP3 target	PI target	LEAFY target	SVP target	AG target	AGL15 target
AT3G19184	*REM1*	–	–	–	Yes	–	–	–
AT5G60140	*REM11*	–	–	–	–	–	Yes	–
AT3G46770	*REM13*	–	–	–	–	–	–	Yes
AT5G57720	*REM15*	–	Yes	Yes	–	–	Yes	Yes
AT4G34400	*REM17*	–	–	–	Yes	–	–	Yes
AT3G06160	*REM18*	–	–	–	Yes	–	–	–
AT3G06220	*REM19*	–	Yes	Yes	–	–	Yes	–
AT5G18000	*REM20*	–	–	Yes	–	–	Yes	Yes
AT2G24680	*REM30*	–	–	–	–	Yes	–	–
AT2G24700	*REM32*	Yes	–	Yes	–	–	–	–
AT4G31610	*REM34*	Yes	Yes	Yes	–	–	–	–
AT4G31615	*REM35*	Yes	Yes	Yes	–	Yes	–	–
AT4G31620	*REM36*	Yes	Yes	Yes	–	–	–	Yes
AT4G31630	*REM37*	Yes	Yes	Yes	–	–	–	–
AT4G31640	*REM38*	–	Yes	–	–	–	–	–

FIG. 5. High confidence AP1-, SVP-, PI- and AP3-binding sites in the cluster of *REM34* homologues on chromosome 4.

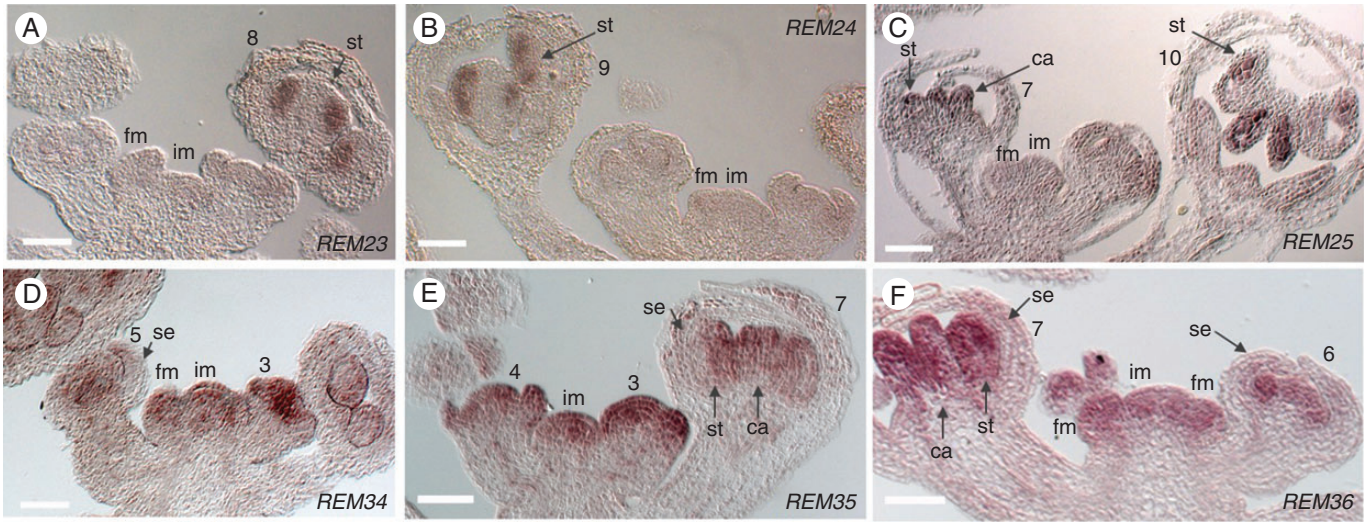

FIG. 6. Expression analysis by *in situ* hybridization for selected *REM* genes in wild-type Col inflorescences. *REM23* (A), *REM24* (B) and *REM25* (C) are expressed from stage 7–8 of flower development. *REM23* and *REM24* expression is detectable in developing stamens (st), whereas *REM25* is expressed from stage 7 of flower development in developing carpels (ca) and stamens. *REM34* (D), *REM35* (E) and *REM36* (F) are all expressed in the inflorescence meristem (im) and floral meristems at stages 1 (fm), 2, 3 and 4. During later stages of flower development, the expression is detectable in the inner floral whorls and excluded from the sepals (se). Scale bars = 50 μm.

expressed in the inflorescence meristem, floral meristem and at later stages in the inner floral whorls, and being excluded from the developing sepals (Fig. 6D–F).

Functional redundancy within the REM *gene family*

To understand the function of the selected *REM* genes, we have analysed insertion mutants of *REM24* and *REM34* and of their closest homologues, *REM23* and *REM36*. Insertion lines for each gene where analysed but showed no peculiar phenotype in flower development with respect to the wild type. In particular, we detected no variation in floral organ number, identity and structure, and inflorescence phyllotaxis. We analysed the *rem24* mutant in which the mRNA is 2·6-fold downregulated, the *rem23* mutant, which is a complete knock-out, the *rem34* mutant, in which the mRNA is reduced by 2·56-fold, and the *rem36* mutant, in which the mRNA is 2·3-fold downregulated (Table 2; Supplementary Data Fig. S3). Of the mutants that we analysed, only *rem23* is a complete knock-out. We decided to analyse different mutant combinations by crossing *rem24* and *rem34* because of their co-expression with the FMI genes *AP1* and *LFY*, and *rem23* with *rem24* because of their similarity in expression profile and their similarity in sequence. No peculiar phenotypes were detected in the *rem24 rem34* and *rem23 rem24* double mutants. No suitable mutant lines were found for *REM25*.

TABLE 2. *Details for the insertion lines that were analysed*

Insertional line	Gene affected	T-DNA in	Effect
SALK_054142	*rem24*	3′UTR	mRNA downregulated 2·6-fold
FLAG_089C09	*rem23*	Second exon	Knock-out
FLAG_566H04	*rem34*	Fourth intron	mRNA downregulated 2·56-fold
SAIL_116500	*rem36*	Putative 5′UTR	mRNA downregulated 2·3-fold

UTR, untranslated region.

DISCUSSION

The fact that the *REM* gene family has been studied for many years (Romanel *et al.*, 2009, 2011; Wynn *et al.*, 2011) without much progress in understanding their roles during plant development highlights the difficulties in discovering *REM* gene functions. Functional redundancy is common among plant transcription factors (Briggs *et al.*, 2006; Hauser *et al.*, 2013) and might also be concealing the function of REM transcription factors.

We have shown that *REM* genes are preferentially expressed during flower and ovule/seed development. However, for some *REM* genes, expression was also detected in vegetative tissues

(Fig. 2). If *REM* genes are highly redundant and active preferentially during flower development, then one might expect to observe effects on plant development when the few *REM* genes that are active in other tissues are knocked-out. This hypothesis is confirmed for *VRN1/REM5* that is also expressed during the floral transition and has been shown to play an important role in flowering time control. *VRN1/REM5* is also expressed at later stages of flower development, but here it might become functionally redundant with other *REM* genes (Levy *et al*., 2002).

To obtain a better understanding of the role that *REM* genes might play during flower development, we have performed a co-expression analysis (Menges *et al*., 2007, 2008). Such an analysis is highly significant since it is calculated using almost 2000 microarray hybridizations. The data set we used (available on NASCarray) includes microarray analyses from different developmental phases, mutants, hormonal treatments, environmental changes, etc. The observed co-expression of *REM* genes with well-studied regulators of flower development can be considered as a strong indication of a functional relationship to these genes since the same procedure has already been validated for other genes (for a review, see Usadel *et al*., 2009). Furthermore, integrating these data with the meta-analysis of available ChIP-seq data and with *in situ* hybridization on the inflorescence apex and early stages of flower development (Figs 5 and 6) strongly confirms that *REM* genes could play a role during early stages of flower development. Therefore, our analysis provides an excellent starting point for future studies regarding REM functions.

Using all the data obtained, we suggest that the *REM34*, *REM35*, *REM36* group is the most interesting one since: (1) *REM34* is co-expressed with the FMI genes; (2) they are bound by AP1, SVP, AP3 and PI; and (3) they are expressed in the floral meristem and during the earliest stages of flower development. *REM34*, *REM35* and *REM36* are closely related homologues and they are clustered on chromosome 4 within <10 kbp. Single mutants of *REM34* and *REM35* apparently do not show any difference in flower development, and thus we hypothesized that such genes may be highly redundant. Unfortunately they are in close linkage and we could not analyse multiple mutant combinations. In the future it would be interesting to knock-down all of these genes using an RNA interference approach (Abbott *et al*., 2002) or to produce multiple knock-outs using genome editing technologies (Miller *et al*., 2011; Cong *et al*., 2013; Jiang *et al*., 2013).

Our analyses highlight the difficulties in studying this gene family due to redundancy and genomic positions, despite the huge amount of information that is available nowadays in different databases as well as the enormous quantity of data arising from high-throughput studies, which all together clearly suggest that this family should be important for reproductive development.

SUPPLEMENTARY DATA

Supplementary data are available online at www.aob.oxfordjournals.org and consist of the following. Table S1: list of oligonucleotides used in this work. Table S2: correlation values among the *REM* gene family that were used to construct Fig. 3 (see legend) are listed in worksheet 'REM correl Lin'. The other two worksheets represent the correlation values calculated from expression values as such ('Full correl table Lin') or after log transformation ('Full correl table Log') for the REM family, the ARF family and several genes involved in meristem and flower development. The conditional formatting highlights cells whose values are comprised between 0·5 and 0·6, between 0·6 and 0·75 and $>0{\cdot}75$ in grey, yellow and red background, respectively. Figure S1: co-expression matrix for REMs and key regulators of reproductive development without clustering. Figure S2: expression levels of *REM* genes measured by qPCR. Figure S3: expression levels of *REM* genes measured by qPCR in wild-type plants and in the respective mutants for *REM23*, *REM24*, *REM34* and *REM36*.

ACKNOWLEDGEMENTS

We thank Luca Mizzi for his help with the correlation analyses and their representation; and Giulia Leo, Caterina Selva, Beatrice Castelnovo and Luis Mathias-Hernandez for technical support. The PhD fellowship of O.M. and the post-doctoral fellowship of V.G. were supported by the Università degli Studi di Milano. M.A.M. was supported by the Marie Curie European Union Initial Training Network (ITN) SYSFLO project. M.A.M., M.A.F., C.M.P., S.M.N. and L.C. were supported by IRSES-GA-2009-247587 – EVOCODE.

LITERATURE CITED

Abbott JC, Barakate A, Pincon G, Legrand M, Lapierre C, Mila I, Schuch W, Halpin C. 2002. Simultaneous suppression of multiple genes by single transgenes. Down-regulation of three unrelated lignin biosynthetic genes in tobacco. *Plant Physiology* **128**: 844–853.

Beekwilder J, van Leeuwen W, van Dam NM, *et al*. 2008. The impact of the absence of aliphatic glucosinolates on insect herbivory in arabidopsis. *PLoS One* **3**: e2068.

Berri S, Abbruscato P, Faivre-Rampant O, *et al*. 2009. Characterization of WRKY co-regulatory networks in rice and arabidopsis. *BMC Plant Biology* **9**: 120.

Briggs GC, Osmont KS, Shindo C, Sibout R, Hardtke CS. 2006. Unequal genetic redundancies in arabidopsis – a neglected phenomenon? *Trends in Plant Science* **11**: 492–498.

Coen ES, Romero JM, Doyle S, Elliott R, Murphy G, Carpenter R. 1990. *floricaula*: homeotic gene required for flower development in *Antirrhinum majus*. *Cell* **63**: 1311–1322.

Cong L, Ran FA, Cox D, *et al*. 2013. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**: 819–823.

Ezcurra I, Wycliffe P, Nehlin L, Ellerström M, Rask L. 2000. Transactivation of the Brassica napus napin promoter by ABI3 requires interaction of the conserved B2 and B3 domains of ABI3 with different cis-elements: B2 mediates activation through an ABRE, whereas B3 interacts with an RY/G-box. *The Plant Journal* **24**: 57–66.

Favaro R, Pinyopich A, Battaglia R, *et al*. 2003. MADS-box protein complexes control carpel and ovule development in arabidopsis. *The Plant Cell* **15**: 2603–2611.

Franco-Zorrilla JM, Fernández-Calvín B, Madueño F, Cruz-Alvarez M, Salinas J, Martínez-Zapater JM. 1999. Identification of genes specifically expressed in cauliflower reproductive meristems. Molecular characterization of BoREM1. *Plant Molecular Biology* **39**: 427–436.

Franco-Zorilla JM, Cubas P, Jarillo J, Fernàndez-Calvìn B, Salinas J, Martınez-Zapater JM. 2002. AtREM1, a member of a new family of B3 domain-containing genes, is preferentially expressed in reproductive meristems. *Plant Physiology* **128**: 418–427.

Gregis V, Sessa A, Colombo L, Kater MM. 2008. AGAMOUS-LIKE24 and SHORT VEGETATIVE PHASE determine floral meristem identity in arabidopsis. *The Plant Journal* **56**: 891–902.

Gregis V, Sessa A, Dorca-Fornell C, Kater MM. 2009. The arabidopsis floral meristem identity genes AP1, AGL24 and SVP directly repress class B and C floral homeotic genes. *The Plant Journal* **60**: 626–637.

Gregis V, Andrés F, Sessa A, *et al.* 2013. Identification of pathways directly regulated by SHORT VEGETATIVE PHASE during vegetative and reproductive development in arabidopsis. *Genome Biology* **14**: R56.

Hauser F, Chen W, Deinlein U, *et al.* 2013. A genomic-scale artificial microRNA library as a tool to investigate the functionally redundant gene space in arabidopsis. *The Plant Cell* **25**: 2848–2863.

Irish VF, Sussex M. 1990. Function of the apetala-1 gene during Arabídopsis floral development. *The Plant Cell* **2**: 741–753.

Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP. 2013. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Research* **41**: e188.

Kagaya Y, Ohmiya K, Hattori T. 1999. RAV1, a novel DNA-binding protein, binds to bipartite recognition sequence through two distinct DNA-binding domains uniquely found in higher plants. *Nucleic Acids Research* **27**: 470–478.

Kaufmann K, Muiño JM, Jauregui R, *et al.* 2009. Target genes of the MADS transcription factor SEPALLATA3: integration of developmental and hormonal pathways in the arabidopsis flower. *PLoS Biology* **7**: e1000090.

Kaufmann K, Wellmer F, Muiño JM, *et al.* 2010. Orchestration of floral initiation by APETALA1. *Science* **328**: 85–89.

King GJ, Chanson AH, McCallum EJ, *et al.* 2013. The arabidopsis B3 domain protein VERNALIZATION1 (VRN1) is involved in processes essential for development, with structural and mutational studies revealing its DNA-binding surface. *Journal of Biological Chemistry* **288**: 3198–3207.

Krizek BA, Fletcher JC. 2005. Molecular mechanisms of flower development: an armchair guide. *Nature Reviews Genetics* **6**: 688–698.

Levy YY, Mesnage S, Mylne JS, Gendall AR, Dean C. 2002. Multiple roles of arabidopsis VRN1 in vernalization and flowering time control. *Science* **297**: 243–246.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**: 402–408.

Matias-Hernandez L, Battaglia R, Galbiati F, *et al.* 2010. VERDANDI is a direct target of the MADS domain ovule identity complex and affects embryo sac differentiation in arabidopsis. *The Plant Cell* **22**: 1702–1715.

McCarty DR, Hattori T, Carson CB, Vasil V, Lazar M, Vasil IK. 1991. The Viviparous-1 developmental gene of maize encodes a novel transcriptional activator. *Cell* **66**: 895–905.

Mendes MA, Guerra RF, Berns MC, *et al.* 2013. MADS domain transcription factors mediate short-range DNA looping that is essential for target gene expression in arabidopsis. *The Plant Cell* **25**: 2560–2572.

Menges M, Pavesi G, Morandini P, Bögre L, Murray JAH. 2007. Genomic organization and evolutionary conservation of plant D-type cyclins. *Plant Physiology* **145**: 1558–1576.

Menges M, Dóczi R, Okrész L, *et al.* 2008. Comprehensive gene expression atlas for the arabidopsis MAP kinase signalling pathways. *New Phytologist* **179**: 643–662.

Miller JC, Tan S, Qiao G, *et al.* 2011. A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology* **29**: 143–148.

Murgia I, Tarantino D, Soave C, Morandini P. 2011. Arabidopsis CYP82C4 expression is dependent on Fe availability and circadian rhythm, and correlates with genes involved in the early Fe deficiency response. *Plant Physiology* **168**: 894–902.

Mylne JS, Barrett L, Tessadori F, *et al.* 2006. LHP1, the arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC. *Proceedings of the National Academy of Sciences, USA* **103**: 5012–5017.

Ó'Maoiléidigh DS, Wuest SE, Rae L, *et al.* 2013. Control of reproductive floral organ identity specification in arabidopsis by the C function regulator AGAMOUS. *The Plant Cell* **25**: 2482–2503.

Ratcliffe OJ, Bradley DJ, Coen ES. 1999. Separation of shoot and floral identity in arabidopsis. *Development* **126**: 1109–1120.

Romanel EA, Schrago CG, Couñago RM, Russo CAM, Alves-Ferreira M. 2009. Evolution of the B3 DNA binding superfamily: new insights into REM family gene diversification. *PLoS One* **4**: e5791.

Romanel E, Das P, Amasino RM, Traas J, Meyerowitz E, Alves-Ferreira M. 2011. Reproductive Meristem22 is a unique marker for the early stages of stamen development. *International Journal of Developmental Biology* **55**: 657–664.

Schmid M, Davison TS, Henz SR, *et al.* 2005. A gene expression map of Arabidopsis thaliana development. *Nature Genetics* **37**: 501–506.

Suzuki M, Kao CY, McCarty DR. 1997. The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity. *The Plant Cell* **9**: 799–807.

Swaminathan K, Peterson K, Jack T. 2008. The plant B3 superfamily. *Trends in Plant Science* **13**: 647–655.

Ulmasov T, Hagen G, Guilfoyle T. 1997. ARF1, a transcription factor that binds to auxin response elements. *Science* **276**: 1865–1868.

Usadel B, Obayashi T, Mutwil M, *et al.* 2009. Co-expression tools for plant biology: opportunities for hypothesis generation and caveats. *Plant, Cell and Environment* **32**: 1633–1651.

Wang Y, Barbacioru C, Hyland F, *et al.* 2006. Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. *BMC Genomics* **7**: 59.

Wang Y, Deng D, Zhang R, Wang S, Bian Y, Yin Z. 2012. Systematic analysis of plant-specific B3 domain-containing proteins based on the genome resources of 11 sequenced species. *Molecular Biology Reports* **39**: 6267–6282.

Weigel D, Alvarez J, Smyth DR, Yanofsky MF, Meyerowitz EM. 1992. LEAFY controls floral meristem identity in arabidopsis. *Cell* **69**: 843–859.

Wellmer F, Alves-Ferreira M, Dubois A, Riechmann JL, Meyerowitz EM. 2006. Genome-wide analysis of gene expression during early arabidopsis flower development. *PLoS Genetics* **2**: e117.

Winter CM, Austin RS, Blanvillain-Baufumé S, *et al.* 2011. LEAFY target genes reveal floral regulatory logic, cis motifs, and a link to biotic stimulus response. *Developmental Cell* **20**: 430–443.

Wuest SE, Maoileidigh DSO, Rae L, *et al.* 2012. Molecular basis for the specification of floral organs by APETALA3 and PISTILLATA. *Proceedings of the National Academy of Sciences, USA* **109**: 13452–13457.

Wynn AN, Rueschhoff EE, Franks RG. 2011. Transcriptomic characterization of a synergistic genetic interaction during carpel margin meristem development in Arabidopsis thaliana. *PLoS One* **6**: e26231.

Yant L, Mathieu J, Dinh TT, *et al.* 2010. Orchestration of the floral transition and floral development in arabidopsis by the bifunctional transcription factor APETALA2. *The Plant Cell* **22**: 2156–2170.

Zhao S, Fernald RD. 2005. Comprehensive algorithm for quantitative real-time polymerase chain reaction. *Journal of Computational Biology* **12**: 1047–1064.

Zheng Y, Ren N, Wang H, Stromberg AJ, Perry SE. 2009. Global identification of targets of the arabidopsis MADS domain protein AGAMOUS-Like15. *The Plant Cell* **21**: 2563–2577.