BIOLOGIA PLANTARUM 46 (3): 321-330, 2003

REVIEW

New roles for MADS-box genes in higher plants

F. GARCÍA-MAROTO*[1], M.-J. CARMONA**, J.-A. GARRIDO*, M. VILCHES-FERRÓN***, J. RODRÍGUEZ-RUIZ*** and D. LÓPEZ ALONSO***

Departamento de Bioquímica, Departamento de Biología Aplicada**, Universidad de Almería, E-04120 Almería, Spain*
*Departamento de Biotecnología, Universidad Politécnica de Madrid, E-28040 Madrid, Spain****

Abstract

Putative transcription factors bearing a particular DNA-binding domain called "MADS-box", have been mainly involved in processes related to flower development. It is generally accepted that MADS-box genes may have played a central role in the evolution of plant reproductive structures. During the last years increasing evidence points to more general roles of these factors that spans to the control of the flowering time, but also to other non-reproductive processes. Moreover, sequencing of the *Arabidopsis* genome has led to the recognition of above hundred MADS-box genes in this model organism, most of them still uncharacterized. This opens the possibility of uncovering new roles for MADS-box genes in plant development and evolution.

Additional key words: *Arabidopsis*, flowering time, plant development, transcription factor, vegetative development.

Introduction

The "MADS" term gathers the initials from the *Minichromosomal maintenance 1* (*MCM1*), *Agamous* (*AG*), *Deficiens* (*DEF*) and *Serum response factor* (*SRF*) genes, all of them sharing the so-called MADS-box (Schwartz-Sommer *et al.* 1990). MADS-box genes make up a multigene family encoding putative transcription factors which are characterized by a conserved DNA-binding/dimerization domain (Shore and Sharrocks 1995, Riechmann and Meyerowitz 1997). In addition to their wide distribution among animals, plants and fungi, some homology has also been found in bacteria, supporting the notion that the MADS-box could already be present in the common ancestor to prokaryotes and eukaryotes (Mushegian and Koonin 1996). During the last decade a huge amount of genetic and molecular information has accumulated, mainly from plant models like *Arabidopsis*, *Antirrhinum* and *Petunia*, leading to the notion that MADS-box genes could have played an important role in the evolution of the reproductive structures in plants. This has opened the way to investigate the evolutionary origin of the morphological diversity among plant species (for recent reviews see Irish 2000, Theissen *et al.* 2000, Ng and Yanofsky 2001). Although most of the available knowledge refers to the function of MADS-box genes in the differentiation of floral/inflorescence and floral organ meristems, growing evidence has pointed out to the involvement of this gene family also in the vegetative development of the plant. In this review we focus on the role of plant MADS-box genes in processes different to those of flower development.

Received 16 January 2002, *accepted* 9 May 2002.
Abbreviations: AGL - agamous like; GA - gibberellins; LD - long day; MIKC - MADS-box intermediate keratine-like C-terminal domain structure; MADS - minichromosomal maintenance 1 agamous-deficiens serum response factor; NMIKC - N-terminal MIKC domain structure; SD - short day.
[1]Author for correspondence; fax: (+34) 950 015008, e-mail: fgmaroto@ual.es

Structure and evolution of MADS factors

Most of plant MADS proteins initially described shared a similar structure designated as MIKC, according to the initials of their characteristic domains (Fig. 1*B*). The MADS-box domain, placed at the N-terminus, is composed by a sequence of about 60 amino acids pretty well conserved among different MADS-box genes and organisms. It folds to give an structural motif with an antiparallel coiled coil of two α-helices (one from each monomer) that is involved in the specific binding to DNA sequences (CArG boxes) conforming the consensus $CC(A/T)_6GG$ (Pellegrini *et al.* 1995). A second conserved domain, the K-box spans about 70 amino acids and is characterized by regularly spaced hydrophobic residues. The resulting structure is an amphipathic α-helix (keratin-like) that is responsible for protein dimerization. This particular domain seems to be an evolutionary acquirement exclusive of plant MADS factors, since it is absent from fungal or animal MADS proteins (Theissen *et al.* 2000, Alvarez-Buylla *et al.* 2000b). The intermediate (I) or linker (L) domain is a short variable segment of about 30 residues located between the K and MADS boxes. Early studies have shown that MADS proteins bind DNA as dimers, and that the molecular determinants controlling dimerization specificity are distributed along the MADS, I and K domains, varying among different MADS proteins (Riechmann *et al.* 1996). Recent interaction models propose that DNA binding could be achieved through the formation of multimeric complexes of MADS factors that would interact with a pair of CArG boxes brought into proximity through DNA bending (Theissen and Saedler 2001, Theissen 2001). Finally, the C-terminal domain is a sequence and length-variable amino-acid stretch that is thought to mediate protein-protein interactions acting as a transactivation domain. In addition, some MADS-box factors possess an additional N-terminal extension, N-domain, and are accordingly designated as NMIKC.

Recent phylogenetic analysis in *Arabidopsis thaliana* has led to the proposal of two evolutionary lineages represented by Type I and Type II MADS-box genes (Alvarez-Buylla *et al.* 2000b). They differ both in the amino acid sequence of the MADS-box as well as in the domain structure of the predicted protein. The MADS-box of Type I factors seems to have a higher homology to that of animal and fungi SRF-like factors, while for Type II proteins the MADS-box sequence is clearly more related to that of MEF2 (myocyte enhancer factor 2) like factors, also represented in animals and fungi. In addition,

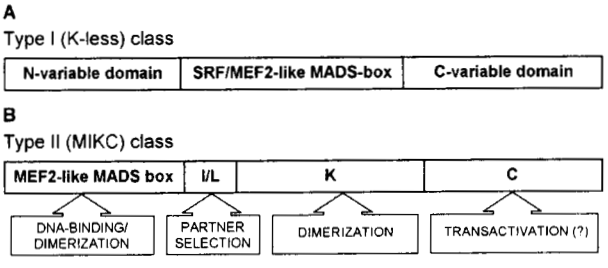

Fig. 1. Structure and domain function of plant MADS-box factors. Phylogenetic analysis and domain structure has lead to the proposal of two MADS-box classes or lineages in plants. A. Type I genes bear a MADS-box with a high homology to that of SRF-like factors and lack the K-box that is substituted by a C-terminal variable domain. B. Type II genes exhibit an MEF2-like MADS-box and the typical MIKC structure. Some of them also possess a variable N-terminal extension, not drawn in the scheme. Domain functions demonstrated or predicted for Type II factors are shown below the corresponding boxes.

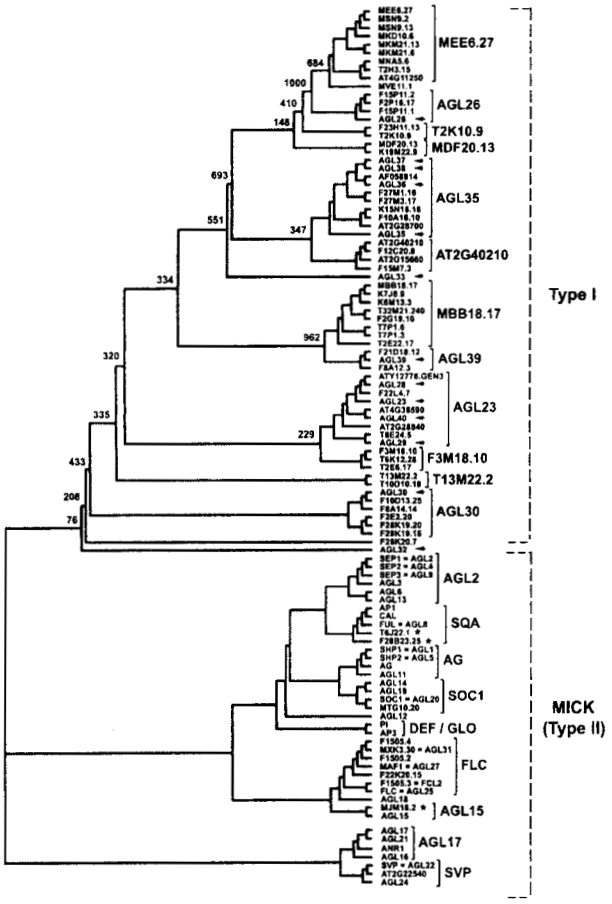

Fig. 2. Phylogenetic analysis of the MADS-box family in *Arabidopsis*. A total of 104 protein sequences comprising the MADS-box plus some 20 - 50 additional residues were aligned using the *Clustal X* (v.1.8) program using the default settings. The alignment containing 120 positions was used to construct the unrooted neighbor-joining tree. Clusters for Type II factors were designated as in (Theissen *et al.* 1996), or arbitrarily, for the Type I class. *Arrows* indicate those Type I factors already cited in literature. Type II factors not previously reported are marked by an *asterisk*. Bootstrap values (over 1000 replicates) are indicated for relevant nodes of the Type I class. MADS-box proteins not previously reported were named according to their gene designation as stated in the *Arabidopsis* databank.

Type I factors lack the K-box, conforming a structure with a MADS-box followed by a rather undefined and length-variable domain, while most of Type II factors exhibit the typical MIKC structure. In addition, an N-terminal variable extension is found in a number of both Type I and Type II factors. Nevertheless, monophyly of the Type I lineage was not well defined, as a number of genes lacking the K-box are also present in *Arabidopsis* whose MADS-box can be either clearly assigned to the MEF2 type, or difficult to classify in any of the SRF or MEF2 groups. The study was based on the available sequence information at that time (45 MADS-box genes). Upon complete sequencing of the *Arabidopsis* genome it is now possible to get a whole picture of the MADS-box family. In this work we present a whole phylogenetic analysis (Fig. 2), where MADS box proteins deduced from genes found in the *Arabidopsis* database, a total of 104, have been included. 38 of them correspond to MIKC/Type II genes, characterized by the classical MIKC structure and the presence of a MEF2-like MADS-box. These genes appear grouped in accordance to the clusters previously described, with the inclusion of three new members, not previously reported, placed in clusters SQUAMOSA (SQUA) and AGL15. The 66 remaining genes group together in a big cluster that include 12 *agamous*-like genes (*AGLs*) previously designated as Type I or unclassified genes (Alvarez-Buylla *et al.* 2000b). Although statistical support (bootstrap values) is, as in the previous report, not good for this clade, we will keep the Type-I designation of this group for simplicity. Within this group, an enormous heterogeneity exists related to both domain structure and MADS-box sequence, although they all lack the K-box. Thus, while the MADS-box of genes within the AGL35 cluster seem to have a higher homology to SRF-like MADS-box (based on the conservation of a number of characteristic residues), others, like those in the AGL23 or T13M22.2 clades, remain difficult to assign to any of the two groups. The question whether a SRF-like MADS-box is present in plant factors is relevant because it would indicate, as it has been proposed, that the evolutionary split between SRF and MEF2-like MADS-box genes took place before separation between animal/fungi and plant kingdoms (Alvarez-Buylla *et al.* 2000b). On the other hand, it is evident that the reduced number of the previously reported Type I genes just represented the iceberg tip of a quantitatively important group that contains even more members than the better known Type II class. This favours the notion that plant genes lacking the K-box may represent active genes and not merely evolutionary by-products (*i.e.* pseudogenes). Moreover, the almost absolute lack of information about these genes opens a wide investigation field that may uncover the participation of MADS-box genes in new cellular processes.

Functions performed by MADS-box genes

A great diversity of functions have been described for MADS-box genes, covering both developmental and metabolic processes. Thus, the yeast ARG80 factor is involved in the control of arginine metabolism (Dubois *et al.* 1987), while MCM1 regulates transcription of different genes involved in cell cycle, growth and cell differentiation (Passmore *et al.* 1988). In insects, the MADS-box genes so far known have been involved in the development of the muscle and trachea (Affolter *et al.* 1994, Lilly *et al.* 1994). For vertebrates, described functions include muscle development (MEF-like factors) and mediation of the cellular response to growth factors (SRF). The first two MADS-box genes described in plants in early nineties were *deficiens* from *Antirrhinum majus* and *agamous* from *Arabidopsis thaliana,* both involved in flower development. Since then, about 40 genes have been uncovered in *Arabidopsis,* a number that is probably less than half of the whole family size in this organism (Fig. 2, Riechmann *et al.* 2000), which is by now the main reference in higher plants. Most of the known functions are related to the flowering process either by controlling the appropriate development of floral meristem (meristem identity genes), or the different flower organs (organ identity genes). The analysis of mutants and the expression patterns of genes involved have led to the proposal of genetic models explaining gene interactions during flower development (reviewed in Theissen 2001). In addition to this, recent studies have demonstrated the involvement of MADS-box genes in the control of flowering time, fruit development, as well as in different aspects of the vegetative development.

MADS-box genes and the control of flowering time

The switch from vegetative to reproductive development is regulated by a complex genetic network that integrates both endogenous and environmental signals, mainly corresponding to the age of the plant, photoperiodic conditions and temperature. The analysis of mutants affected in flowering time, mainly in *Arabidopsis*, as well as molecular data from a number of the involved genes, have revealed the existence of several genetic pathways

leading to flowering (Fig. 3), for which a high redundancy exists (for review see Koornneef *et al.* 1998, Simpson *et al.* 1999, Reeves and Coupland 2000, Araki 2001). The following pathways have been described in *Arabidopsis*, although their characteristics may differ in other plant species depending on the particular photoperiodic/vernalization requirements.

The so-called autonomous pathway monitors somehow the developmental state (age) of the plant by sensing endogenous signals. The vernalization pathway allows the plant to flower early when exposed to low temperatures. Light quality and day-length signals are channelled through the so-called photoperiodic pathway which acts by accelerating flowering under long-day photoperiods (LD) relative to short days (SD). A fourth pathway has been proposed in which gibberellin phytohormones (GA) would promote flowering through a day-length independent pathway that is essential under short day conditions. A number of MADS-box genes have been shown to participate within this network (Table 1, Fig. 3).

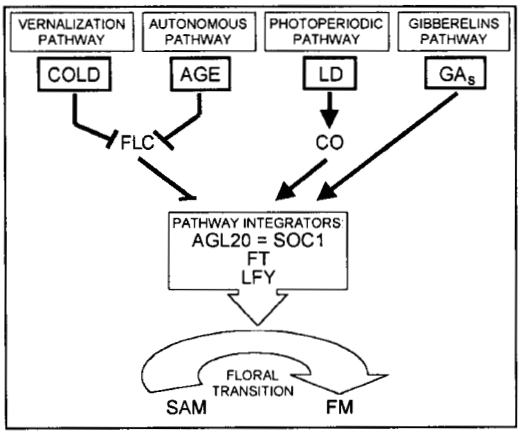

Fig. 3. Flowering pathways in *Arabidopsis*. Genetic analysis support the existence of four genetic pathways leading to flowering in *Arabidopsis* (see explanation in the text). A number of so-called pathway integrator genes receive inputs from the different pathways leading to the floral transition from the vegetative shoot apical meristem (SAM) to the reproductive floral / inflorescence meristem (FM).

Table 1. *Arabidopsis* MADS-box genes involved in the control of flowering time.

Name	Effect on flowering transition	Flowering pathway	Other functions	Putative orthologues (organism)
FLC/FLF (*AGL25*)	repressor	autonomous and vernalization	unknown	unknown
MAF1/FLM (*AGL27*)	repressor	autonomous and vernalization	unknown	unknown
SOC1 (*AGL20*)	promoter	autonomous, vernalization and photoperiodic	unknown	*SAMADSA* (*Sinapis alba*) *MADS1* (*Nicotiana tabacum*)
FUL (*AGL8*)	promoter	unknown	carpel and fruit development; floral identity	*SAMADSB* (*Sinapis alba*)
SVP (*AGL22*)	repressor	autonomous (?)	unknown	unknown
AGL24	promoter (?)	unknown	unknown	*STMADS16* (*Solanum tuberosum*)

Genetic analysis of the plant response to vernalization allowed the identification of the locus *flowering locus C*. The corresponding gene (*FLC*) has been cloned independently by T-DNA tagging (Sheldon *et al.* 1999) and chromosome walking (Michaels and Amasino 1999), and shown to encode a MADS-box protein previously designated as AGL25. This factor acts as a dosage-dependent repressor of flowering as demonstrated by the fact that an increased gene activity causes a strong delay in flowering while an early-flowering phenotype is obtained by antisense reduction of the transcript content (Michaels and Amasino 1999, Sheldon *et al.* 1999, 2000). Those studies have shown that *FLC* is mainly regulated through inputs from the autonomous and the vernalization pathways (Fig. 3). Thus, *FLC* mRNA content is reduced by a cold treatment that lead to flowering acceleration. A negative effect is also achieved by the autonomous pathway as relatively high transcript amounts are detected in mutants of this pathway characterized by late-flowering phenotypes. *FLC* is generally expressed along the plant, even in vegetative tissues. However, no obvious alterations of the vegetative organs are observed in loss-of-function mutants. It is also remarkable that expression of *FLC* in the whole plant remains constant throughout development, something that is explained by considering that down-regulation of the repressor must take place in a limited number of cells within the apical meristem. Nevertheless, *in situ* expression analysis should give a clear answer to that question. Regarding *FLC* targets, it has been proposed that it negatively regulates two key flowering-time genes, *FT* and *SOC1*, (Fig. 3) that act as integrators of signals from different flowering pathways (Lee *et al.* 2000, Samach *et al.* 2000). Little is known about the molecular mechanism by which the *FLC* transcriptional activity is regulated. Interestingly, a reduction of the DNA methylation in the cell leads to a reduction of the *FLC* transcript content and, consequently, to early flowering, thus paralleling the vernalization effect (Finnegan *et al.* 1998, Sheldon *et al.* 1999). Although this correlation

implies that *FLC* is regulated by the methylation status, it remains unclear whether the methylation effect is achieved directly on the FLC gene or on other *FLC* regulators. On the other hand, a relationship between *FLC* and the gibberellin pathway has been proposed indicating that FLC may act counteracting the promotion of flowering by GA in the apical meristem (Sheldon *et al.* 1999).

Recently, a MADS-box gene, *MADS affecting flowering 1* (*MAF1*), previously designated as *AGL27*, has been reported in *Arabidopsis* that behaves similarly to *FLC* (Ratcliffe *et al.* 2001). It has been suggested that *MAF1* acts either downstream or independently of *FLC*. *MAF1* exhibits a generalized expression pattern in the plant. The existence of multiple MAF1/AGL27 transcripts has been reported, possibly generated through alternative splicing (Sheldon *et al.* 1999, Alvarez-Buylla *et al.* 2000a). This is interesting as structurally different MADS proteins are predicted to be generated that could perform different functions in either the same or different tissues of the plant. Therefore, the study of the expression pattern for the individual transcripts is of paramount importance for an appropriate interpretation of results such as those obtained from "sense" and "antisense" analysis. In agreement with participating in the same process, *MAF1* lies in the *FLC* cluster (Fig. 2). This cluster also includes five still uncharacterized MADS-box genes whose functions could also be related to the control of the flowering time (Ratcliffe *et al.* 2001). The identification of the product of the *flowering locus M* has been achieved (Scortecci *et al.* 2001), and shown to be identical to MAF1.

Another MADS-box gene recently identified as a flowering promoter, integrating signals from the different pathways is the *supressor of overexpression of constans* (*SOC1*) or *AGL20*. This has been independently shown through different experimental approaches. Late-flowering mutants affected in *AGL20* have been identified by transposon tagging (Borner *et al.* 2000), thus indicating an involvement of this gene in the control of the flowering time. On the other hand, *AGL20*, later called *SOC1*, has been identified as a direct target of *constans* (*CO*), a gene involved in the photoperiodic pathway (Samach *et al.* 2000). *AGL20* is also regulated by genes belonging to the autonomous pathway, as demonstrated by a down-regulation of this gene in mutants affected in both the photoperiodic and the autonomous pathways. In addition, regulation of *AGL20* by gibberellins has been shown (Borner *et al.* 2000). According to this, *SOC1/AGL20* has been proposed to play a central role as an integrator of the signals from the different flowering pathways. Expression of *AGL20* in *Arabidopsis* shows again a generalized pattern that is affected by the photoperiod. A low content of the *AGL20* transcript is found in the apical meristem when plants are grown under SD, while a considerable and fast increase is obtained when they are induced to flowering by shifting to LD. Up-regulation is also obtained during the transition to flowering under SD, something that is consistent with the proposed regulation by signals from the different pathways. *AGL20* is also expressed at later stages in the inflorescence meristem, and later on, in the centre of the flower meristem. This would indicate a possible role in flower development, although *AGL20/SOC1* mutations apparently do not affect flower morphology. An *AGL20* orthologous gene has been identified in the related species *Sinapis alba* (Bonhomme *et al.* 2000, Borner *et al.* 2000). Interestingly, over-expression of this gene in a SD cultivar of *Nicotiana tabacum* was able to overcome the absolute requirement of SD for flowering, thus circunventing the photoperiodic block imposed under LD (Borner *et al.* 2000). Moreover, activation of this gene by gibberellins and cytokinins is observed, according to the notion that AGL20 activation is one of the steps of the gibberellin/cytokinin pathway leading to flowering (Bonhomme *et al.* 2000).

Fruitful (*FUL*, former *AGL8*) is another MADS-box gene that has been involved in the control of the floral transition. Early studies on this gene revealed that it plays a clear role during carpel and fruit development (Gu *et al.* 1998), as it will be discussed in a next section. In addition, it has been shown that *FUL* also participates in the acquirement of the floral meristem identity together and redundantly with *apetala 1 (AP1)* and *cauliflower (CAL)*. Although the *ful* mutant is only slightly affected in flowering time, triple mutants for *FUL*, *CAL* and *AP1* causes a non-flowering phenotype, thus indicating its role in promoting floral transition (Ferrándiz *et al.* 2000a). According to this function, *FUL* is expressed weakly in rossette leaves during vegetative development and strongly upregulated in the shoot apex upon transition to flowering (Gu *et al.* 1998). In addition, *FUL* (together with *AP1* and *CAL*) acts, in turn, as a regulator of *leafy* (*LFY*) and *terminal flower 1* (*TFL1*), two other meristem identity genes (Ferrándiz *et al.* 2000a). Consistent with the similarity of functions, *FUL* lies within the same cluster than *AP1* and *CAL*. It is remarkable the finding during our search of two new genes, T6J22.1 and F28B23.25, belonging to this cluster and not previously reported in *Arabidopsis* (Fig. 2). It is likely that their function could also be related to the control of the flowering time and/or the specification of meristem identity.

The *short vegetative phase* (*SVP or AGL22*) is another MADS-box gene identified through transposon tagging as a dose-dependent repressor of flowering (Hartmann *et al.* 2000). Loss-of-function mutants for *SVP* are characterised by early-flowering phenotypes under any photoperiodic condition. Expression analysis indicates the existence of two different transcripts with a predominantly vegetative pattern. As one would predict for a flowering repressor, expression in the apical meristem is strongly reduced after floral transition. Transcript levels are not affected by the photoperiodic

conditions or vernalization, thus indicating a possible location within the autonomous pathway. SVP shares a high homology with AGL24, making up together a separate genetic cluster (Fig. 2). It has been suggested that *AGL24* acts as a promoter of the flowering process (Ferrándiz *et al.* 2000a), although this question still awaits further investigation. In this context, a putative *AGL24* orthologue has been identified in potato (*STMADS16*) with a similar expression pattern (García-Maroto *et al.* 2000). Over-expression of *STMADS16* in a tobacco heterologous system accelerates flowering under SD but not in LD thus suggesting a possible role in promoting flowering and a limiting function under SD.

Functional redundancy is a reiterative theme in MADS-box genes and flower development. Given the overlapping expression pattern of *FUL* and *AGL20* in the apical meristem during floral transition, it remains an open question the possibility that *AGL20* and *FUL*, and perhaps other genes as *AGL24*, act redundantly in promoting floral transition as it has been proposed (Ferrándiz *et al.* 2000a). In most cases, placing of MADS-box genes within the genetic pathways leading to flowering remains to be elucidated, and their number will probably increase with the identification of new MADS-box genes.

MADS-box genes in fruit development

Fruit development is a poorly known process, even though one can envisage that its understanding may have a considerable interest, as a way to improve certain fruit qualities. In *Arabidopsis*, three MADS-box genes have been involved up to now. *FUL*, a gene also regulating floral transition and meristem identity, participates in the development of the fruit valves, while *shatterproff 1* (*SHP1*) and *shatterproof 2* (*SHP2*) are essential for the differentiation of the valve dehiscence zone. Fruit dehiscence and, consequently, seed dispersal in *Arabidopsis* and other plants bearing siliques is achieved through differentiation of a few cell layers at the valve margin. Upon fruit maturation a dehiscence zone develops allowing detachment of the valves and releasing of the seeds. *FUL* function in fruit development was uncovered through the analysis of the transposon-generated mutant *fruitful* (*ful-1*). In *ful-1* plants elongation of the silique after fertilization is abolished due to the failure of valve cells to elongate and differentiate, leading to crowded seeds and dehiscence failure (Gu *et al.* 1998). *FUL* exhibits a complex expression pattern (see previous section) that includes the carpel walls, consistently with its role in valve development. Loss-of-function mutants of *SHP1* and *SHP2* have been generated by T-DNA insertion that allowed functional studies (Liljegren *et al.* 2000). Even though single mutations do not show an apparent effect on fruit development, in the double mutant the dehiscence zone fails to develop thus indicating that both genes act redundantly to specify cell-fate within the valve margin. Expression of *SHP1* and *SHP2* takes place, in addition to the valve margin, in other tissues as in developing ovules, septum, nectaries and style. Nevertheless, development of these organs is not apparently affected in the double mutant perhaps due to functional redundancy with other genes in those cells. *FUL* and *SHP* genes play different specialized roles during fruit development and their expression is regulated in an antagonistic way. Thus, *FUL* acts as a negative regulator of *SHP1/2* ensuring that their expression and, consequently, fruit abscission takes place exclusively at the valve margin (Ferrándiz *et al.* 2000b). On the other hand, upstream positive regulation of SHP1/2 is achieved by the product of *AG* (Savidge *et al.* 1995), a MADS-box gene involved in flower development.

Another interesting aspect of these group of genes comes from their phylogenetic relationships. *SHP1* and *SHP2* (previously designated as *AGL1* and *AGL5*) belong to a cluster of genes that include to *AG* among other genes (Fig. 2). Members of this group share similar expression patterns and functions, mainly related to stamen and carpel development. On the other hand, *FUL* belongs to a cluster integrated by genes like *AP1* and *CAL* whose functions are related with the specification of flower/inflorescence meristem identity. Although *FUL* still conserve this kind of function, its participation in fruit development may be interpreted as another example of functional recruitment in the course of evolution (Theissen *et al.* 2000).

Investigation in plants other than *Arabidopsis* still remains very scarce. Recently, a novel gene, *DEFH28*, from the snapdragon *Antirrhinum majus* has been described that is likely to represent the ortholog of *FUL* (Muller *et al.* 2001). A dual role both in the control of floral meristem identity and fruit development was also shown for *DEFH28*. Another gene called *jointless* has been characterized in *Lycopersicon esculentum* that seems to be essential in the process leading to pedicel abscission. *Jointless* was isolated by chromosome walking, and further shown to be responsible for the *jointless* mutation by co-segregation analysis (Mao *et al.* 2000). *Jointless* mutants are characterized among other features by the failure to develop the abscission zone that normally allows shedding of the fruit. In this sense *jointless* function resembles that of *SHP* genes although at a different spatial location in the fruit. However, assignation of *JOINTLESS* to phylogenetic clusters puts *JOINTLESS* in a different clade, together with *AGL24* and *SVP*, two flowering time genes. In agreement to this, *jointless* mutants are also affected in the regulation of the

meristem fate, with inflorescence meristems reverting to vegetative growth after forming few flowers (Szymkowiak and Irish 1999). As in the case of their clade partners, *JOINTLESS* also shows a generalized expression pattern, with the exception of the green tissue of the fruit calix.

Initial studies of B-function genes like *deficiens* and *globosa* from *Antirrhinum majus*, which are involved in the determination of flower organ identity, revealed that expression of these genes was maintained throughout the development of the flower and that it also included carpel tissues. Consequently, it was postulated that they could play additional roles during fruit development (Tröbner *et al.* 1992, García-Maroto *et al.* 1993). This turned out to be true as it has been recently reported the involvement of the *globosa* orthologous gene from *Malus domestica* (*MdPI*) in the parthenocarpic development of the fruit (Yao *et al.* 2001), thus uncovering a new function for these MADS-box sub-family. This is of biotechnological importance as it opens the possibility for the generation of high-quality seedless cultivars through down-regulation of this gene.

Table 2. MADS-box genes involved in fruit development.

Gene	Organism	Function in fruit development	Other functions
FUL (AGL8)	*Arabidopsis*	valve development	flowering time/floral meristem identity/leaf development
DEFH28	*Antirrhinum*	carpel wall development	floral meristem identity
SHP1(AGL1)	*Arabidopsis*	valve margin abscission	unknown
SHP2(AGL5)	*Arabidopsis*	valve margin abscission	unknown
JOINTLESS	*Lycopersicum esculentum*	pedicel abscission	inflorescence/ floral meristem identity
MDPI	*Malus domestica*	parthenocarpic development	flower organ identity

MADS-box genes in non-reproductive processes

Some of these genes were likely to participate in the development of other plant structures, as they are expressed in vegetative tissues such as leaves, stem and roots (Ma *et al.* 1991). Although there is a general correlation between expression pattern and place of function, one should keep in mind that the presence of transcript in a given tissue would not necessarily means that gene activity is required for its proper development. We will first described those cases in which direct evidence exists for the involvement of MADS-box genes in vegetative processes, and finally some genes for which such a role might be inferred from their expression patterns.

Root development: The MADS-box gene *ANR1* was identified in *Arabidopsis* as a result of a screening for nitrate inducible genes (Zhang and Forde 1998). Expression of *ANR1* is restricted to the roots of plants and is positively regulated by nitrate. This behaviour seems to be part of a response of the plant to the presence of nitrate in the soil, a situation in which lateral root growth is stimulated. Knocking-out of *ANR1* by antisense technology eliminates this response thus indicating that *ANR1* is required for lateral root development. It has been proposed that *ANR1* could act as a transcriptional regulator of genes modulating the rate of lateral root elongation. Phylogenetic analysis shows that *ANR1* is located in a cluster that includes to *AGL17*, a gene that is also specifically expressed in roots, among other genes like *AGL16* and *AGL21*, for which little information is still available (Fig. 2).

Root nodulation: Another MADS gene, *NMH7*, has been reported in *Medicago sativa* whose expression is specifically induced in root nodules upon infection by *Rhizobium* (Heard and Dunn 1995). Interestingly, homology search for a putative orthologue of *NMH7* in the *Arabidopsis* genome does not give any clear

Table 3. MADS-box genes involved in vegetative development.

Gene	Organism	*Arabidopsis* orthologue	Expression pattern	Function
ANR1	*Arabidopsis*	---	Root	Lateral root development
NMH7	*Medicago sativa*	Unknown	Nodule	Nodule differentiation (?)
NMH5	*Medicago sativa*	AGL17	Root/Nodule	Unknown
AGL17	*Arabidopsis*	---	Root	Unknown
FUL	*Arabidopsis*	---	Generalized	Leaf development and others

candidate. The most similar is *apetala 3* (90 % identity at the MADS-box, 54 % overall identity), a B-function gene classically involved in the development of flower sepals and stamens. Thus, *NMH7* could represent another case of MADS-box gene recruitment, in this case from the cluster of B-function genes, to perform some specialized function at least in legume plants. Moreover, an *AGL17* orthologous gene (*NMH5*) has been described in alfalfa whose transcripts are also found in nodules (Heard *et al.* 1997), although its function in these symbiotic organs still remains obscure. In this context, it is interesting to notice that the development of root nodules is a response of the legume plant to a low soil nitrogen content. As *ANR1* is up-regulated by the nitrate, it is possible that the activation of the closely related gene, *NMH5*, was induced by the nitrate accumulation within the root-derived nodule cells.

Leaf development: Although a MADS-box gene specifically involved in leaf development has not been described yet there is evidence that *FUL* activity is required for leaf cellular differentiation. This is inferred from the fact that cauline leaves from the *ful-1* mutant are broader that those in the wild type, together with a reduction in the number of internal cell layers (Gu *et al.* 1998). Association of *FUL* expression to the vascular system of vegetative tissues argues in favour of an effect on vein differentiation as the cause of changes in the leaf shape, although this interpretation still awaits further studies.

Inferred functions from expression patterns: For a number of MADS-box genes the only available information is the expression pattern, although some inferences have been done about their roles. Among them, other genes from *Arabidopsis* whose transcripts are exclusively found in root cells, and that could be related to some aspect of root development are *AGL14* and *AGL19*, both included in the *TM3* clade, and the solitary *AGL12* (Rounsley *et al.* 1995, Alvarez-Buylla *et al.* 2000a). Other MADS-box genes have been proposed to be involved in seed development. Detailed expression studies have been performed with the seed specific gene *AGL15* (Heck *et al.* 1995, Rounsley *et al.* 1995, Perry *et al.* 1996). AGL15 protein is found very early in embryo development and maintained at a high level until initiation of seed dessication. AGL15 seems to be also present in the endosperm tissues of dicot and monocot species (Perry *et al.* 1996), although some discrepancy exists since in some studies the corresponding transcript was not detected in the endosperm (Rounsley *et al.* 1995). In addition, nuclear localization of AGL15 and binding of this protein to a CArG motif has been shown, indicating a role of AGL15 as a transcriptional regulator. Search in the *Arabidopsis* genome database reveals the existence of an *AGL15* related gene designated MJM18.2, not previously reported. In the cluster analysis (Fig. 2) both genes appear grouped, and it is therefore conceivable that they are involved in the same developmental process.

AGL18 is another MADS-box gene whose transcripts are found in the seed (Alvarez-Buylla *et al.* 2000a). Nevertheless, this gene has a complex expression pattern that includes both vegetative and reproductive tissues, being remarkable its expression during pollen development. Within the seed, *AGL18* mRNA is specifically found in the endosperm but not in embryonic cells, a characteristic that makes it a useful seed marker for this kind of tissue.

AGL16 is another MADS-box gene belonging to the *AGL17* cluster (Fig. 2). However, its expression is not restricted to the roots, exhibiting a generalized expression in the plant (Alvarez-Buylla *et al.* 2000a). Interestingly, expression in the leaf, where the mRNA is found at a higher level, is mainly associated to stomatal guard cells and trichomes, while in the roots it is again found in the epidermal cells. Consequently, a possible a role in epidermal differentiation processes giving raise to those structures has been suggested. It is likely that future studies will reveal the participation of MADS-box factors like AGL16 in cell-type specification networks that are being uncovered in leaves and roots (Larkin *et al.* 1997, Schiefelbein *et al.* 1997).

Conclusions

The MADS-box gene family encode transcriptions factors that perform different functions during plant development. Among their roles are not only the development of flower organs, but also the control of the flowering time and the development of organs such as fruits, roots, leaves, and other specialized tissues. Investigation in this field is still very scarce for most of plants. Even in the plant model *Arabidopsis* the number of MADS-box factors characterized so far is much below the whole number uncovered by the sequencing of its genome. This makes likely that new functional aspects of MADS-box genes are still revealed as knowledge increase in the next years.

References

Affolter, M., Montagne, J., Walldorf, U., Groppe, J., Kloter, U., LaRosa, M., Gehring, W.J.: The *Drosophila* SRF homolog is expressed in a subset of tracheal cells and maps within a genomic region required for tracheal development. - Development **120**: 743-753, 1994.

Alvarez-Buylla, E.R., Liljegren, S.J., Pelaz, S., Gold, S.E., Burgeff, C., Ditta, G.S., Vergara-Silva, F., Yanofsky, M.F.: MADS-box gene evolution beyond flowers: expression in pollen, endosperm, guard cells, roots and trichomes. - Plant J. **24**: 457-466, 2000a.

Alvarez-Buylla, E.R., Pelaz, S., Liljegren, S.J., Gold, S.E., Burgeff, C., Ditta, G.S., Pouplana, L.R., Martinez-Castilla, L., Yanofsky, M.F.: An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. - Proc. nat. Acad. Sci. USA **97**: 5328-5333, 2000b.

Araki, T.: Transition from vegetative to reproductive phase. - Curr. Opin. Plant Biol. **4**: 63-68, 2001.

Bonhomme, F., Kurz, B., Melzer, S., Bernier, G., Jacqmard, A.: Cytokinin and gibberellin activate *SaMADS A*, a gene apparently involved in regulation of the floral transition in *Sinapis alba*. - Plant J. **24**: 103-111, 2000.

Borner, R., Kampmann, G., Chandler, J., Gleissner, R., Wisman, E., Apel, K., Melzer, S.: A MADS domain gene involved in the transition to flowering in *Arabidopsis*. - Plant J. **24**: 591-599, 2000.

Dubois, E., Bercy, J., Messenguy, F.: Characterization of two genes, *ARGRI* and *ARGRIII*, required for specific regulation of arginine metabolism in yeast. - Mol. gen. Genet. **207**: 142-148, 1987.

Ferrándiz, C., Gu, Q., Martienssen, R., Yanofsky, M.F.: Redundant regulation of meristem identity and plant architecture by *FRUITFULL*, *APETALA1* and *CAULIFLOWER*. - Development **127**: 725-734, 2000a.

Ferrándiz, C., Liljegren, S.L., Yanofsky, M.F.: Negative regulation of the *SHATTERPROOF* genes by *FRUITFULL* during *Arabidopsis* fruit development. - Science **289**: 436-438, 2000b.

Finnegan, E.J., Genger, R.K., Kovac, K., Peacock, W.J., Dennis, E.S.: DNA methylation and the promotion of flowering by vernalization. - Proc. nat. Acad. Sci. USA **95**: 5824- 5829, 1998.

García-Maroto, F., Ortega, N., Lozano, R., Carmona, M.-J.: Characterization of the potato MADS-box gene *STMADS16* and expression analysis in tobacco transgenic plants. - Plant mol. Biol. **42**: 499-513, 2000.

García-Maroto, F., Salamini, F., Röhde, W.: Molecular cloning and expression patterns of three alleles of the *Deficiens*-homologous gene *St-Deficiens* from *Solanum tuberosum*. - Plant J. **4**: 771-780, 1993.

Gu, Q., Ferrándiz, C., Yanofsky, M.F., Martienssen, R.: The *FRUITFULL* MADS-box gene mediates cell differentiation during *Arabidopsis* fruit development. - Development **125**: 1509-1517, 1998.

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

Heard, J., Caspi, M., Dunn, K.: Evolutionary diversity of symbiotically induced nodule MADS box genes: characterization of nmhC5, a member of a novel subfamily. - Mol. Plant Microbe Interact. **10**: 665-676, 1997.

Heard, J., Dunn, K.: Symbiotic induction of a MADS-box gene during development of alfalfa root nodules. - Proc. nat. Acad. Sci. USA **92**: 5273-5277, 1995.

Heck, G.R., Perry, S.E., Nichols, K.W., Fernandez, D.E.: AGL15, a MADS domain protein expressed in developing embryos. - Plant Cell **7**: 1271-1282, 1995.

Irish, V.F.: Variations on a theme: flower development and evolution. - Genome Biol. Rev. **1**: 1015.1-1015.4, 2000.

Koornneef, M., Alonso-Blanco, C., Peeters, A.M.J., Soppe, W.: Genetic control of flowering time in *Arabidopsis*. - Annu. Rev. Plant Physiol. Plant mol. Biol. **49**: 345-370, 1998.

Larkin, J.C., Marks, M.D., Nadeau, J., Sack, F.: Epidermal cell fate and patterning in leaves. - Plant Cell **9**: 1109-1120, 1997.

Lee, H., Suh, S.-S., Park, E., Cho, E., Ahn, J.H., Kim, S.-G., Lee, J.S., Kwon, Y.M., Lee, I.: The AGAMOUS-like 20 MADS domain protein integrates floral inductive pathways in *Arabidopsis*. - Genes Dev. **14**: 2366-2376, 2000.

Liljegren, S.L., Ditta, G.S., Eshed, Y., Savidge, B., Bowman, J.L., Yanofsky, M.F.: *SHATTERPROOF* MADS-box genes control seed dispersal in *Arabidopsis*. - Nature **404**: 766-770, 2000.

Lilly, B., Galewsky, S., Firulli, A.B., Schulz, R.A., Olson, E.N.: a MADS box transcription factor expressed in differentiating mesoderm and muscle cell lineages during *Drosophila* embryogenesis. - Proc. nat. Acad. Sci. USA **91**: 5662-5666, 1994.

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

Mao, L., Begum, D., Chuang, H.-W., Budiman, M.A., Szymkowlak, E.J., Irish, E.E., Wing, R.A.: *JOINTLESS* is a MADS-box gene controlling tomato flower abscission zone development. - Nature **406**: 910-913, 2000.

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

Muller, B.M., Saedler, H., Zachgo, S.: The MADS-box gene DEFH28 from *Antirrhinum* is involved in the regulation of floral meristem identity and fruit development. - Plant J. **28**: 169-179, 2001.

Mushegian, A.R., Koonin, E.V.: Sequence analysis of eukaryotic developmental proteins: ancient and novel domains. - Genetics **144**: 817-828, 1996.

Ng, M., Yanofsky, M.F.: Function and evolution of the plant MADS-box gene family. - Nature Rev. **2**: 186-195, 2001.

Passmore, S., Maine, G.T., Elble, R., Christ, C., Tye, B.-K.: A *Saccharomyces cerevisiae* protein involved in plasmid maintenance is necessary for mating of MATa cells. - J. mol. Biol. **204**: 593-606, 1988.

Pellegrini, L., Tan, S., Richmond, T.J.: Structure of serum response factor core bound to DNA. - Nature **376**: 490-498, 1995.

Perry, S.E., Nichols, K.W., Fernandez, D.: The MADS domain protein AGL15 localizes to the nucleus during early stages of seed development. - Plant Cell **8**: 1977.1989, 1996.

Ratcliffe, O.J., Nadzan, G.C., Reuber, T.L., Riechmann, J.L.: Regulation of flowering in *Arabidopsis* by an FLC homologue. - Plant Physiol. **126**: 122-132, 2001.

Reeves, P.H., Coupland, G.: Response of plant development to

environmental control of flowering by daylength and temperature. - Curr. Opin. Plant Biol. **3**: 37-42, 2000.

Riechmann, J.L., Krizek, B.A., Meyerowitz, E.M.: Dimerization specificity of *Arabidopsis* MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. - Proc. nat. Acad. Sci. USA **93**: 4793-4798, 1996.

Riechmann, J.L., Meyerowitz, E.M.: MADS domain proteins in plant development. - Biol. Chem. **378**: 1079-1101, 1997.

Riechmann, J.L., Heard, J., Martin, G., Reuber, L., Jiang, C.-Z., Keddie, J., Adam, L., Pineda, O., Ratcliffe, O.J., Samaha, R.R., Creelman, R., Pilgrim, M., Broun, P., Zhang, J.Z., Ghandehari, B.K., Sherman, B.K., Yu, G.-L.: *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. - Science **290**: 2105-2110, 2000.

Rounsley, S.D., Ditta, G.S., Yanofsky, M.F.: Diverse roles for MADS box genes in *Arabidopsis* development. - Plant Cell **7**: 1259-1269, 1995.

Samach, A., Onouchi, H., Gold, S.E., Ditta, G.S., Schwarz-Sommer, Z., Yanofsky, M.F., Coupland, G.: Distinct roles of CONSTANS target genes in reproductive development of *Arabidopsis*. - Science **288**: 1613-1616, 2000.

Savidge, B., Rounsley, S.D., Yanofsky, M.F.: Temporal relationship between the transcription of two *Arabidopsis* MADS box genes and the floral organ identity genes. - Plant Cell **7**: 721-733, 1995.

Scortecci, K.C., Michaels, S.D., Amasino, R.M.: Identification of a MADS-box gene, FLOWERING LOCUS M, that represses flowering. - Plant J. **26**: 229-236, 2001.

Schiefelbein, J.W., Massuci, J.D., Wang, H.: Building a root: the control of patterning and morphogenesis during root development. - Plant Cell **8**: 1089-1098, 1997.

Schwartz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H., Sommer, H.: Genetic control of flower development by homeotic genes in *Antirrhinum majus*. - Science **250**: 931-936, 1990.

Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J., Dennis, E.S.: The *FLF* MADS box gene: A repressor of flowering in *Arabidopsis* regulated by vernalization and methylation. - Plant Cell **11**: 445-458, 1999.

Sheldon, C.C., Rouse, D.T., Finnegan, J., Peacock, W.J., Dennis, E.S.: The molecular basis of vernalization: The central role of *FLOWERING LOCUS C* (*FLC*). - Proc. nat. Acad. Sci. USA **97**: 3753-3758, 2000.

Shore, P., Sharrocks, A.D.: The MADS-box family of transcription factors. - Eur. J. Biochem. **229**: 1-13, 1995.

Simpson, G.G., Gendall, A.R., Dean, C.: When to switch to flowering. - Annu. Rev. cell dev. Biol. Plant **15**: 519-550, 1999.

Szymkowiak, E.J., Irish, E.E.: Interactions between *jointless* and wild-type tomato tissues during development of the pedicel abscission zone and the inflorescence meristem. - Plant Cell **11**: 159-175, 1999.

Theissen, G.: Development of floral organ identity: stories from the MADS house. - Curr. Opin. Plant Biol. **4**: 75-85, 2001.

Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J.T., Münster, T., Winter, K.-U., Saedler, H.: A short history of MADS-box genes in plants. - Plant mol. Biol. **42**: 115-149, 2000.

Theissen, G., Kim, J.T., Saedler, H.: Classification and phylogeny of the MADS-Box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes. - J. mol. Evol. **43**: 484-516, 1996.

Theissen, G., Saedler, H.: Floral quartets. - Nature **409**: 469-471, 2001.

Tröbner, W., Ramierez, L., Motte, P., Hue, I., Huijser, P., Lönnig, W.E., Saedler, H., Sommer, H., Schwarz-Sommer, Z.: *GLOBOSA*: a homeotic gene which interacts with *DEFICIENS* in the control of *Antirrhinum* floral organogenesis. - EMBO J. **11**: 4693-4693-4704, 1992.

Yao, J.-L., Dong, Y.-H., Morris, B.A.M.: Parthenocarpic apple fruit production conferred by transposon insertion mutations in a MADS-box transcription factor. - Proc. nat. Acad. Sci. USA **98**: 1306-1311, 2001.

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