MADS-box Genes in Plant Evolution and Development

Barbara A. Ambrose*

The New York Botanical Garden, 200th St. & Southern Blvd., Bronx, NY 10458, USA

Correspondence: * bambrose@nybg.org

ABSTRACT

Evolutionary developmental studies have shown that large transcription factor families underlie key morphological features. The duplication and diversification of these transcription factor families has been important to provide material for the generation of new interacting partners where duplicates could take on a new role or a sub-function of the original protein. Arguably one of the most important of these transcription factor families in plants is the MADS-box family of transcription factors. MADS-box genes are an ancient group found in animals, plants and fungi but have duplicated and diversified much more in plants than in animals and fungi. The plant MADS-box genes can be further sub-divided into Type I, Type II MIKCc and Type II MIKC* MADS-box-genes. The Type II MIKCc MADS-box genes are best known for their role in the ABC model of floral organ identity. MIKCc MADS-box genes are also important for flowering time, fruit, endosperm and seed development. There have been many Type II MIKCc studies performed across the land plants however, comparatively little is known about the Type II MIKC* and Type I MADS-box genes. Therefore, recent functional analyses of the previously uncharacterized MIKC* and Type I MADS-box genes are particularly exciting. Recent genome analyses have shown that MIKC* and Type I MADS-box genes can also be found in *Physcomitrella patens* and *Selaginella moellendorffii*. Phylogenetic, functional and expression analyses of MADS-box genes across the land plants will provide insight into the role of all types of MADS-box genes in the evolution and development of land plant body plans.

CONTENTS

INTRODUCTION

Large families of transcription factors have been the focus of evolutionary developmental (evo-devo) studies. In animals, the duplication and loss of members of the large homeobox (HOX) transcription factor family have been implicated in the evolution and development of diverse animal body plans (Swalla 2006). The animal regulatory protein paired box gene 6 (PAX6) has been co-opted multiple times during eye evolution to specify a light gathering cell that has then been elaborated into multiple different types of eyes (Gehring 2005). These studies in eye evolution and development also illustrate that complex organs such as eyes do not arise de novo but are built using existing building blocks that are co-opted for new or modified functions.

In addition, transcription factors are usually part of a large regulatory network with each protein having several interacting partners (Davidson *et al.* 2002). These proteins are usually composed of multiple domains. The domains within the protein determine the interactions of these proteins within regulatory networks and the addition or loss of domains can modify the interactions within the regulatory network (Galant and Carroll 2002; Veron *et al.* 2007). The animal MADS-box protein myocyte enhancer factor-2 (MEF2) is part of a core regulatory network with other transcription factors and signaling proteins that directs heart development (Olson 2006). The duplication and diversification of the MEF2 regulatory network has been important for the evolution and development of the heart from a simple pump to a complex 4-chambered muscle. The duplication and div-

Invited Review

ersification of transcription factor families has been important to provide material for the generation of new interacting partners where duplicates could take on a new role or a sub-function of the original protein.

One of the best-known regulatory families underlying plant development is the MADS-box family of transcription factors. This family has been particularly well studied for its role in the ABC model of floral organ specification (Krizek and Fletcher 2005). However, the MADS-box floral organ specification genes make up a small portion of this gene family in flowering plants (Parenicova *et al.* 2003). The complete MADS-box gene family is now known from the genomes of the flowering plants *Arabidopsis thaliana* (Arabidopsi*s*), *Oryza sativa* (rice) and *Populus trichocarpa* (poplar) (AGI 2000; Goff *et al.* 2002; Yu *et al.* 2002; Parenicova *et al.* 2003; Leseberg *et al.* 2006; Tuskan *et al.* 2006; Arora *et al.* 2007). Functional studies in Arabidopsis have shown that MADS-box genes have important roles in development besides floral organ specification (Theissen *et al.* 2000). In addition, the recent completion of the land plant genomes from the bryophyte, *Physcomitrella patens* and the lycophyte, *Selaginella moellendorffii*, has provided insights into the evolution of the MADS-box family of transcription factors and will be integral for future studies on the role of MADS-box genes in land plant evolution and development (Rensing *et al.* 2008).

There have been extensive studies on MADS-box genes in the land plants but these have focused on the MIKCc type of MADS-box gene, particularly the floral organ identity genes in angiosperms. These Type II MIKCc MADS-box genes have been identified across the land plants although little is known of their function outside of the angiosperms (Munster *et al.* 1997; Hasebe *et al.* 1998; Winter *et al.* 1999; Krogan and Ashton 2000; Carlsbecker *et al.* 2003; Jager *et al* 2003; Tanabe *et al.* 2003). I will briefly review the background of these important Type II MIKCc MADS-box genes with particular emphasis on functional characterizations and protein-protein-interactions. These analyses illustrate how duplications and diversifications within the MIKCc group may provide starting material for the elaboration of key morphological features. Recent functional analyses of the previously uncharacterized Type II MIKC* and Type I MADS-box genes are particularly exciting. Recent genome analyses have shown that MIKC* and Type I MADS-box genes can also be found in mosses and lycophytes. Comparative expression analyses, phylogenetic analyses, functional analyses and protein-protein interaction studies will elucidate the function of the MIKC* and Type I MADS-box genes and provide insight into their contribution to land plant evolution and development.

MADS-BOX TRANSCRIPTION FACTORS

Antiquity of the MADS domain

The MADS-box family of transcription factors is an ancient family of proteins found in animals, plants and fungi (Alvarez-Buylla *et al.* 2000). The MADS acronym is derived from the first 4 founding members of this family: **M**ini Chromosome Maintenance from yeast, **A**GAMOUS from Arabidopsis, **D**EFICIENS from *Antirrhinum majus* and **S**erum Response Factor from humans (Norman *et al.* 1988; Passmore *et al.* 1988; Schwarz-Sommer *et al.* 1990; Sommer *et al.* 1990; Yanofsky *et al.* 1990). In animals and fungi, there are no more than 10 MADS-box genes found in any one species (Riano-Pachon *et al.* 2008; Riechmann *et al.* 2000). However in the flowering plant genomes of Arabidopsis, rice and poplar 75–107 MADS-box genes have been identified (AGI 2000; Goff *et al.* 2002; Yu *et al.* 2002; Parenicova *et al.* 2003; Leseberg *et al.* 2006; Tuskan *et al.* 2006; Arora *et al.* 2007). The large number of MADS-box genes found in flowering plants is proposed to have arisen by duplication and diversification of function (Purugganan *et al.* 1995).

MADS-box protein classification

MADS-box proteins from animals, plants and fungi all share a 55-60 amino acid conserved DNA-binding domain called the MADS domain (Norman *et al.* 1988; Passmore *et al.* 1988; Schwarz-Sommer *et al.* 1990; Sommer *et al.* 1990; Yanofsky *et al.* 1990). The MADS domain is typically found at the N terminus of the protein. An ancient duplication in the last common ancestor of these 3 eukaryotic kingdoms is proposed to have given rise to 2 lineages of MADS-box proteins: Type I and Type II (**Fig. 1**) (Alvarez-Buylla *et al.* 2000). Animal, plant and fungal Type I and Type II MADS-box proteins were first distinguished by amino acid synapomorphies in the MADS DNA binding domain. Type I MADS-box proteins have an SRF-like MADS domain while Type II MADS-box proteins have a MEF2-like MADS domain (Shore and Sharrocks 1995). There are no other conserved domains outside of the MADS domain shared between animals, plants and fungi (Alvarez-Buylla *et al.* 2000).

In plants, Type I and Type II MADS-box genes can be further distinguished by the protein domains C terminal to the MADS domain and the intron-exon structure of Type I and Type II genes (**Fig. 1**) (Alvarez-Buylla *et al.* 2000; Kaufmann *et al.* 2005). The plant Type II MADS-box proteins, or MIKC proteins, have a well-defined structure (Alvarez-Buylla *et al.* 2000; Kaufmann *et al.* 2005). In addition to the MADS domain, there is an approximately 70 amino acid K domain recognized by its regularly spaced hydrophobic residues that is proposed to form a coiled-coil structure (Ma *et al.* 1991; Davies and Schwarz-Sommer 1994; Riechmann and Meyerowitz 1997). The MADS and K domains are divided by a less well-conserved intervening region or I region of variable length. At the C terminus of the protein is the aptly named C domain. The K domain is involved in protein-protein interactions while the I region may confer some specificity to the protein-protein interactions. The function of the C region is less well defined however, it has been suggested that this region is important as a trans-activation domain and has been shown to be important for ternary complex formation (Riechmann and Meyerowitz 1997; Egea-Cortines *et al.* 1999). The protein domains C-terminal to the MADS-domain in Type I are less well defined however, shared motifs of unknown function have been recognized in some of these proteins (Parenicova *et al.* 2003). Type I MADS-box genes are composed of one or two exons while Type II MADS-box genes are composed of 5-11 exons (Kaufmann *et al.* 2005).

Phylogenetic and sequence analyses in Arabidopsis and the moss, *Physcomitrella patens* have revealed that the Type I and Type II MADS-box proteins can be further subdivided based on sequence and/or shared motifs (Henschel *et al.* 2002; Parenicova *et al.* 2003). The Type II MADS-box genes can be further subdivided into MIKC classic (MIKCc) and MIKC* based on the intron-exon structure of their coding sequences (Henschel *et al.* 2002; Kofuji *et al.* 2003). The MIKC* MADS-box genes have a longer I domain encoded by multiple exons instead of 1 exon (Henschel *et al.* 2002). MIKCc and MIKC* (also known as Mδ) MADS-box genes can also be distinguished by phylogenetic analyses using the MADS domain alone (Parenicova *et al.* 2003). The Type I MADS domain proteins were divided into 3 groups: Mα, Mβ, and Mγ, based on phylogenetic analyses of the MADS domain and some conserved motifs in the C-terminal region (Parenicova *et al.* 2003).

MIKC CLASSIC MADS-BOX GENES

The role of MIKCc MADS-box genes in flower development

MIKCc MADS-box genes are the best-characterized MADS-box genes in any land plant. and are well known for their role in flower development (Krizek and Fletcher 2005). The ABC model of floral organ identity describes how dif-
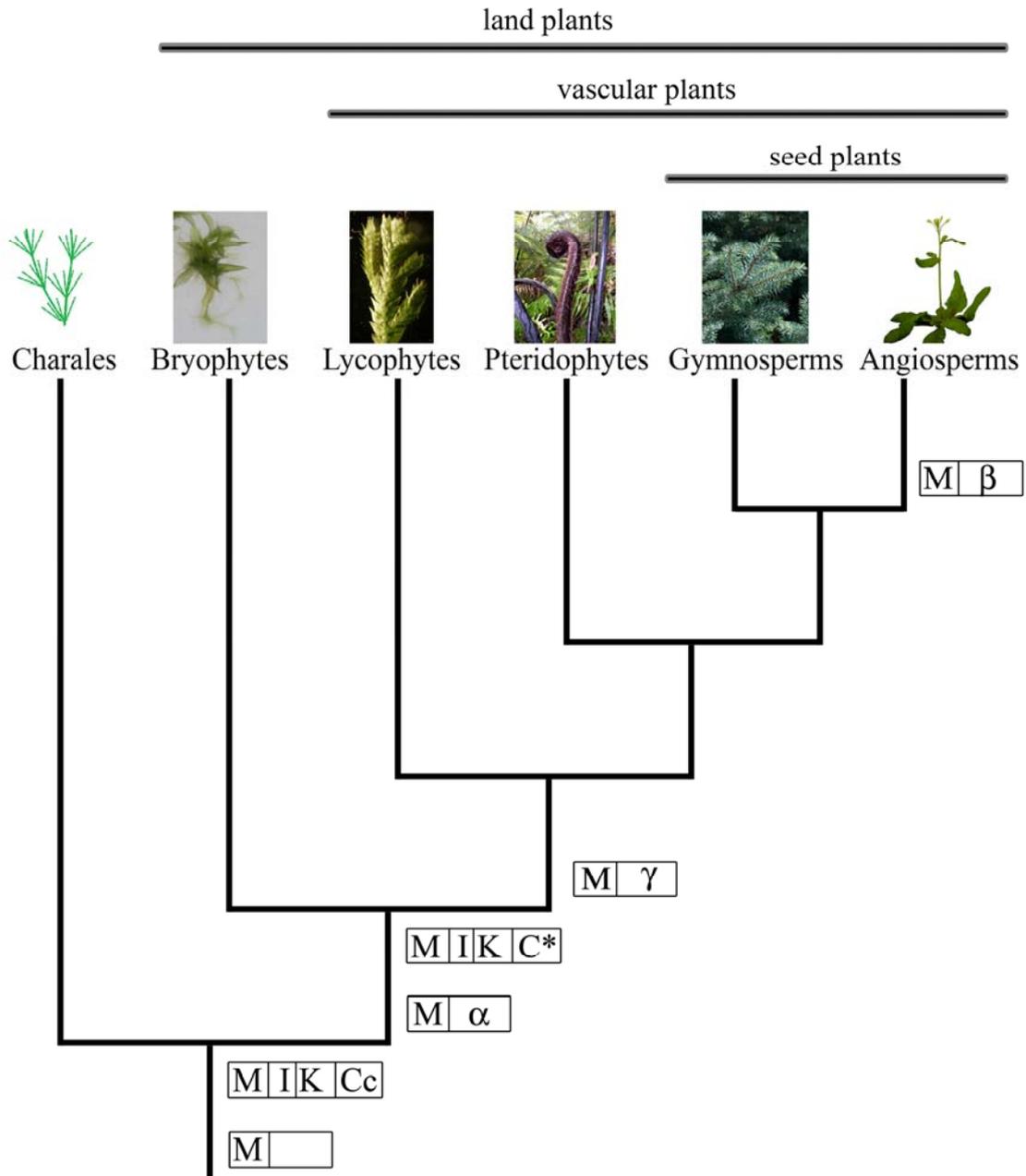
Fig. 1 The evolution of MADS-box proteins in the land plants. The Type I (M) and Type II (MIKC) protein structures are mapped onto a generalized land plant phylogeny. The Type I and Type II MADS-box proteins have similar DNA binding MADS (M) domain but differ in the domains 3′ to the M domain. The C-terminal domain of the Type I proteins is not well defined. The region 3′ to the MADS domain of Type II proteins have 3 recognizable domains: I, K and C. There are 2 kinds of Type II MADS-box proteins: MIKCc and MIKC* and 3 kinds of Type I MADS-box proteins: Mα, Mγ and Mβ. An ancient duplication gave rise to the Type I and Type II lineages so both types of proteins should be found in the green plants although only MIKCc have been isolated from the charophycean green algae. MIKCc MADS-box proteins have been identified throughout the land plants. Genomic and phylogenetic analyses have identified MIKC* in the bryophyte *Physcomitrella patens*, the lycophyte *Lycopodium annotinum* and the angiosperms *Arabidopsis thaliana*, *Oryza sativa* and *Populus trichocarpa* suggesting that MIKC* MADS-box proteins were present in the last common ancestor of bryophytes and vascular plants. Mα MADS-box genes have been identified in the bryophyte *P. patens*, the lycophyte *Selaginella moellendorffii* and the angiosperms *A. thaliana*, *O. sativa* and *P. trichocarpa* also suggesting that Mα MADS-box proteins were present in the last common ancestor of bryophytes and vascular plants. Mγ MADS-box proteins have been identified in the lycophyte *S. moellendorffii* and the angiosperms *A. thaliana*, *O. sativa* and *P. trichocarpa* suggesting that Mγ MADS-box proteins were present in the last common ancestor of vascular plants. Mβ proteins have only been identified in the angiosperms *A. thaliana*, *O. sativa* and *P. trichocarpa*.

ferent combinations of MADS-box proteins are necessary for specifying the 4 floral organ identities sepals, petals, stamens and carpels (Bowman *et al.* 1989; Meyerowitz *et al.* 1991). A loss of A, B or C class gene function results in the homeotic conversion of floral organs and a loss of all ABC functions in Arabidopsis results in a flower composed entirely of leaves. Although there is a large amount of floral diversity in the angiosperms, all flowers have the same basic ground plan of sepals, petals, stamens and carpels. Comparative studies across the angiosperms have shown that the ABC model of floral organ identity is largely con-

served across the 2 major groups of angiosperms: monocots and eudicots (Bowman 1997; Ambrose *et al.* 2000). Network analyses have also shown that the ABC model of floral organ identity is a modular gene regulatory network providing theoretical support for the conservation of the ABC model of floral organ identity (Mendoza *et al.* 1999; Espinosa-Soto *et al.* 2004). Morphological, functional and theoretical analyses on floral organ development suggest that the basic ground plan of the flower is relatively conserved due to developmental drive.

Duplication and diversification of MIKCc clades

MIKCc MADS-box genes have been isolated from charophycean green algae that are sister to all land plants, and ancient duplications and subsequent diversifications have given rise to distinct clades (Karol *et al.* 2001; Tanabe *et al.* 2005; Riano-Pachon *et al.* 2008). Phylogenetic analyses of MIKCc MADS-box indicates that there are 12 clades: A, B, C, E, Bs, SVP, SOC1, FLC, XAL1, ANR1, AGL15 and AGL6 in flowering plants (Purugganan *et al.* 1995; Zhang and Forde 1998; Michaels and Amasino 1999; Winter *et al.* 1999; Hartmann *et al.* 2000; Samach *et al.* 2000; Becker and Theissen 2003; Nam *et al.* 2003; Veron *et al.* 2007; Tapia-Lopez *et al.* 2008). These clades have been given numerous terminologies. The above terminology for the MIKCc clades is based on known functions of proteins within the clade such as the A, B, C and E floral organ identity proteins as well as the founding members of each clade: B-sister (Bs), SHORT VEGETATIVE PHASE (SVP), SUPPRESSOR OF OVEREXPRESSION OF CONSTANS (SOC1), XAANTAL (XAL), and ANR1, AGL6 and AGL15 clades. At least 7 out of these 12 clades existed in the last common ancestor of seed plants (**Fig. 1**) (Becker and Theissen 2003).

Proteins from the A, B, C and E clades function in a combinatorial manner to specify the four floral organs (Purugganan *et al.* 1995; Pelaz *et al.* 2000; Honma and Goto 2001). Proteins from the SVP, FLC, SOC1, AGL15 and XAL1 clades are important for regulating the transition from vegetative to reproductive development (Michaels and Amasino 1999; Hartmann *et al.* 2000; Samach *et al.* 2000; Adamczyk *et al.* 2007; Tapia-Lopez *et al.* 2008). Proteins in the SVP clade are also important for flower development (Gregis *et al.* 2006, 2008). AGL15 also plays a role in embryo development (Heck *et al.* 1995). Proteins from the ANR1 clade are important for root development while analyses of AGL6 indicate it has a role in floral patterning (Rounsley *et al.* 1995; Zhang and Forde 1998; Rijpkema *et al.* 2009).

MIKCc MADS-box genes have been isolated from the pteridophyte, *Ceratopteris richardii* and this fern's MIKCc MADS-box genes form distinct clades that do not cluster with any of the seed plant MIKCc MADS-box clades (Munster *et al.* 1997). MIKCc MADS-box genes have also been isolated from the charophycean green algae and phylogenetic analyses indicate that these genes do not cluster with any of the MIKCc MADS-box genes isolated from the lycophyte, *Lycopodium annotinum*, the fern *C. richardii*, and various gymnosperms or angiosperms (Tanabe *et al.* 2005). Phylogenetic analyses of MIKCc MADS-box genes from the completed genome sequences of the moss *Physcomitrella patens* and the lycophyte *Selaginella moellendorffii* may help to elucidate the evolution of the MIKCc MADS-box genes in land plants. In *P. patens*, only 6 out of the 23 MADS-box genes are MIKCc (Rensing *et al.* 2008). More whole genome data from across the land plants and phylogenetic analyses will help to elucidate the evolution of particular MIKCc clades. Understanding the evolution of the A, B, C and E clades may provide insight into the origin of the flower.

The ABC MADS-box protein-protein interaction network is an important developmental module necessary for the specification of the basic floral ground plan (Veron *et al.* 2007). Phylogenetic analyses and comparative protein interaction studies have shown that protein-protein interaction networks of MIKCc MADS-box proteins in general have been important for their evolution (Veron *et al.* 2007). These analyses show that the general increase in the number of MIKCc MADS-box genes in angiosperms has been due to whole genome duplications, the protein-protein interaction network has high inter-clade connectivity and duplicates are retained after whole genome duplications (Veron *et al.* 2007). Veron *et al.* (2007) suggested that these duplicates are retained due to their ability to form higher order complexes with members of diverse clades and this has

facilitated the subsequent diversification of duplicate genes (Veron *et al.* 2007). For example, MADS-box interactions studies in tomato have shown that there has been conservation as well as divergence in the protein interaction network compared to Arabidopsis and petunia networks (Leseberg *et al.* 2008). This diversification in interacting partners is particularly seen in MADS-box genes that underlie specific functions in tomato such as fruit ripening (Vrebalov *et al.* 2002; Leseberg *et al.* 2008).

Duplication and diversification of genes within A and C clades in Arabidopsis have been important for the specification of flower meristem identity, fruit and ovule development. Duplication and diversification within the A clade has given rise to *CAULIFLOWER* (*CAL*) and *FRUITFULL* (*FUL*) in Arabidopsis (Purugganan *et al.* 1995). *CAL* is functionally redundant with the A class gene *APETALA1* (*AP1*) in specifying floral meristem identity (Kempin *et al.* 1995). While FUL has a role in floral meristem identity it also has diversified into a new role in fruit differentiation (Gu *et al.* 1998; Ferrandiz *et al.* 2000). Within the C clade, duplications have given rise to two functionally redundant proteins SHATTERPROOF 1 and 2 (SHP1 and SHP2) that are important for patterning the fruit (Purugganan *et al.* 1995; Liljegren *et al.* 2000). These functional studies in Arabidopsis have illustrated the importance of duplication and diversification of MADS-box proteins in plant evolution and their roles in flower development. Across the angiosperms, duplications within the A, B and C floral organ identity clades have also occurred and have important functional implications (Litt and Irish 2003; Kramer *et al.* 2004; Irish and Litt 2005). Not only the presence of particular clades but their regulatory networks have been important for morphological evolution at least within the angiosperms.

MIKC* MADS-BOX GENES

Antiquity of MIKC* MADS-box genes

MIKC* MADS-box genes were first identified in the moss *Physcomitrella patens* (Henschel *et al.* 2002). Phylogenetic analyses at the time indicated that there was only one other MIKC* MADS-box protein, LAMB1, identified from the lycophyte, *Lycopodium annotinum* (Svensson *et al.* 2000; Henschel *et al.* 2002). Phylogenetic analyses of MADS-box genes from Arabidopsis revealed the presence of 6 MIKC* MADS-box genes (*AGL30, AGL65, AGL66, AGL67, AGL94* and *AGL104*) present in the genome (AGI 2000; Parenicova *et al.* 2003). The identification of MIKC* MADS-box genes in bryophytes, lycophytes and angiosperms indicates that an MIKC* MADS-box gene was present in the last common ancestor of bryophytes and flowering plants (**Fig. 1**).

Role of MIKC* MADS-box genes in plant development

The role of MIKC* MADS-box genes have recently been studied in Arabidopsis. Transcriptome analyses and RT-PCR have shown that MIKC* MADS-box genes are preferentially expressed in Arabidopsis pollen (Kofuji *et al.* 2003; Honys and Twell 2004). MIKC* MADS-box genes bind a particular CArG box motif that is predominantly found in late stage pollen specific genes (Verelst *et al.* 2007). MIKC* mutants have reduced pollen germination that results in reduced fertility (Verelst *et al.* 2007; Adamczyk and Fernandez 2009). AGL104-GUS expression is detected throughout pollen development from the single nuclear stage to the deposition of the sperm cells in the female gametophyte (Reddy and Ambrose, unpublished data). MIKC* MADS-box genes form 2 monophyletic subgroups P (*AGL30, AGL65* and *AGL94*) and S (*AGL66, AGL67* and *AGL104*) (Nam *et al.* 2004; Adamczyk and Fernandez 2009). Yeast-2-hybrid and bimolecular fluorescence complementation protein interaction studies have shown that MIKC* MADS-box genes form heterodimers and

heterodimers are only formed between members of the S and P MIKC* groups in Arabidopsis (de Folter *et al.* 2005; Verelst *et al.* 2007; Adamczyk and Fernandez 2009). Functional analyses have shown that there is functional redundancy within the S and P MIKC* subgroups (Verelst *et al.* 2007; Adamczyk and Fernandez 2009). Recent studies have also shown that these MIKC* MADS-box genes are part of a large pollen regulatory network, that is composed of additional MADS-box genes, particularly Type I MADS-box genes (Verelst *et al.* 2007; Adamczyk and Fernandez 2009). Therefore similar to MIKCc MADS-box gene duplications, MIKC* duplicates maintain their interaction partners. It is not known whether maintenance of S and P MIKC* interactions has facilitated the neo- or sub- functionalization of Arabidopsis MIKC* paralogs.

There are 23 MADS-box genes in the *Physcomitrella patens* genome and astoundingly 12 of these are MIKC* (Rensing *et al.* 2008). RT-PCR demonstrated that 2 MIKC* MADS-box genes, PPM6 and PPM7, are expressed throughout *Physcomitrella patens* development but are more highly expressed in protonemata compared to gametophores and sporophytes (Riese *et al.* 2005). Phylogenetic analyses of Arabidopsis and *P. patens* MADS-box genes are necessary to determine if the *P. patens* MIKC* MADS-box genes cluster with the Arabidopsis S and P MIKC* subgroups and whether there are novel subgroups in *P. patens*. Although there is little information on MIKC* MADS-box genes in *P. patens*, it is enticing to hypothesize on an ancestral function of MIKC* MADS-box genes in land plants. In Arabidopsis, MIKC* MADS-box genes are important for pollen tube germination and growth. In Arabidopsis, the pollen grain is the male gametophyte that germinates and grows by tip growth. Bryophytes, as represented by *P. patens*, have a dominant gametophyte life cycle and some cells of the gametophyte such as the protonemata grow by tip growth (Menand *et al.* 2007). Functional analyses of MIKC* orthologs in *P. patens* will allow us to determine whether MIKC* MADS-box genes had an ancestral function in tip growth, gametophyte development or specifically in male reproductive development.

TYPE I MADS-BOX GENES

Until recently, the only MADS-box genes that had been functionally characterized in any land plant species were the Type II MIKCc MADS-box genes and expression and phylogenetic analyses of Type I MADS-box genes suggested that these were non-functional (De Bodt 2003; Kofuji 2003). Genomic analyses in Arabidopsis indicated a higher birth and death rate of Type I MADS-box genes which means there was rapid duplication of these genes but these duplicates were not maintained (Nam *et al.* 2004). Phylogenetic analyses also suggested that Type I and Type II MADS-box genes have evolved differently (De Bodt *et al.* 2003; Nam *et al.* 2004). Closely related Type I MADS-box genes are clustered on the Arabidopsis chromosomes indicating that these were generated by recent segmental duplications as opposed to whole genome duplications (De Bodt *et al.* 2003; Nam *et al.* 2004; Veron *et al.* 2007).

Type I functional analyses in Arabidopsis

Recently, 5 Type I MADS-box proteins have been functionally characterized: AGL23, AGL28, AGL61 DIANA (DIA), AGL62 and AGL80 (FEM111) (Portereiko *et al.* 2006; Yoo *et al.* 2006; Bemer *et al.* 2008; Colombo *et al.* 2008; Kang *et al.* 2008; Steffen *et al.* 2008). These characterized Type I MADS-box proteins belong to 2 of the 3 sub-groups of Type I MADS-box proteins. AGL23, AGL28, DIA and AGL62 belong to the Mα subgroup while FEM111 belongs to the Mγ sub-group (Parenicova 2003). Over-expression analyses of AGL28 suggest that it promotes flowering while functional analyses of AGL23, DIA, AGL62 and FEM111 show that they all have a role in female gametophyte and/or endosperm development (Portereiko *et al.* 2006; Yoo *et al.*

2006; Bemer *et al.* 2008; Colombo *et al.* 2008; Kang *et al.* 2008; Steffen *et al.* 2008).

AGL23::GUS is expressed in the functional megaspore and maintained until female gametophyte development is complete (Colombo *et al.* 2008). After fertilization, AGL23:: GUS expression is detected in the embryo and endosperm until the torpedo stage of embryogenesis (Colombo *et al.* 2008). Loss-of-function analyses indicate that AGL23 is necessary for the proper progression of megagametogenesis (Colombo *et al.* 2008). FEM111-GFP expression is detected in the central cell and the endosperm before cellularization and in *fem111* the central cell does not develop properly and subsequent development of the endosperm does not occur (Portereiko *et al.* 2006). DIA is also expressed in the central cell and endosperm (Bemer *et al.* 2008; Steffen *et al.* 2008). The *dia* central cell does not develop properly and subsequent endosperm development does not occur (Bemer *et al.* 2008; Steffen *et al.* 2008). Ectopic expression of female gametophyte synergid and antipodal marker genes in *dia* central cells suggest that the *dia* central cell identity is mis-specified (Steffen *et al.* 2008). Protein-protein interaction studies demonstrated that DIA and FEM111 interact (Bemer *et al.* 2008; Steffen *et al.* 2008). AGL62 expression is detected in the developing endosperm until the cellularization stage (Kang *et al.* 2008). In *agl62*, the endosperm cellularizes prematurely suggesting that AGL62 is important for the timing of endosperm cellularization (Kang *et al.* 2008).

The role of Type I MADS-box genes in land plant development

It appears that no Type I mutants were identified previously by forward genetic screens as they are lethal or affect discrete developmental processes (Portereiko *et al.* 2006; Bemer *et al.* 2008; Colombo *et al.* 2008; Kang *et al.* 2008; Steffen *et al.* 2008). Expression analyses suggest that additional Type I MADS-box genes are involved in gametophyte and/or endosperm development. Transcriptome analyses show that of the MADS-box genes preferentially expressed in the endosperm, a majority are Type I MADS-box genes (Day *et al.* 2008). Expression analyses also indicate a role for the Mγ MADS-box gene *PHERES1* (*AGL37*) in endosperm development (Kohler *et al.* 2003). Expression studies have also shown that Type I Mα MADS-box genes, *AGL29* and *AGL97*, are expressed in pollen (male gametophyte) (Adamczyk and Fernandez 2009). More recent expression analyses show that the Type I genes *PHERES1*, *PHERES2*, *AGL35*, *AGL36*, *AGL40*, *AGL62* and *AGL90* are upregulated in incompatible interspecific crosses (Walia *et al.* 2009). These results indicate that the down regulation of these genes is necessary for proper seed development.

Although there are no defined domains C-terminal to the MADS domain of Type I proteins, there are conserved motifs that suggest that these regions could act as transcription activation domains or form coiled-coil structures (De Bodt *et al.* 2003; Parenicova *et al.* 2003). Protein-protein interaction and genetic studies show that Type I Mα and Mγ proteins preferentially form heterodimers (de Folter *et al.* 2005; Bemer *et al.* 2008; Steffen *et al.* 2008). Functional analyses of the domain C terminal to the MADS domain in Type I proteins and additional protein interaction studies will help to determine if Type I proteins are part of large regulatory networks. Comparative functional analyses of Type I MADS-box proteins across the land plants may help to explain why Type I and Type II MADS-box proteins appear to have distinct evolutionary trajectories.

The recent completion of the bryophyte genome, *P. patens*, and the lycophyte, *S. moellendorffii*, has provided some insight into the evolution of Type I MADS-box genes in land plants (**Fig. 1**) (Rensing *et al.* 2008). Preliminary analyses indicate that only the Mα subgroup is present in *P. patens* while the Mα and Mγ subgroups are found in *S. moellendorffii* genome (Ambrose, unpublished data). These preliminary analyses indicate that the Mα Type I MADS-

box genes can be found in the most recent common ancestor of bryophytes and vascular plants while the Mγ can be found in the common ancestor of the vascular plants. Additional genome sequences and comprehensive phylogenetic analyses are needed to assess the evolution of the Type I MADS-box genes across the land plants. Functional analyses in *P. patens* will help to determine the ancestral function of Mα MADS-box genes and interaction and functional studies need to be performed to determine if Mα and Mγ also preferentially form heterodimers in *S. moellendorffii*. In Arabidopsis, Mα and Mγ proteins preferentially interact and are important for female gametophyte and/or endosperm development, Functional analyses will help to elucidate if a Mα - Mγ regulatory module is present in lycophytes and whether this regulatory module was important for subsequent seed evolution.

It is likely that Type I proteins found in *P. patens* and *S. moellendorffii* have been lost during land plant evolution and will have unique functions in these species. Preliminary *in situ* hybridization results show a Type I protein in *P. patens* to be expressed specifically in the archegonia and antheridia, two structures not found in seed plants (Ambrose, unpublished data). Identification and phylogenetic analyses of Type I MADS-box proteins across the land plants will be imperative to understanding their roles in land plant evolution and development.

CONCLUDING REMARKS

Large regulatory networks underlie morphological features and regulatory networks evolve by the loss and gain of protein domains (Veron *et al.* 2007). MADS-box proteins are modular in organization although Type I and Type II MADS-box genes differ in the domains C-terminal to the MADS domain. Type II MADS-box proteins are necessary for the specification of modular structures such as flowers in angiosperms and fruits in Arabidopsis (Krizek and Fletcher 2005). The duplication and subsequent diversification of Type II MADS-box genes is facilitated by the higher order complexes that it forms (Veron *et al.* 2007; Leseberg 2008). Although, Type I MADS-box genes, by definition, do not have a recognizable I or K domain, regions necessary for protein-protein interactions in Type II MADS, there are conserved domains that could form coiled-coil structures (Parenicova *et al.* 2003). Protein-protein interaction studies and genetic studies show that Type I proteins do form heterodimers and these heterodimers are important for their function (de Folter *et al.* 2005; Bemer *et al.* 2008; Steffen *et al.* 2008). Functional analyses indicate that Type I MADS-box genes are also important for specifying another modular structure, the female gametophyte (Friedman and Williams 2003). Phylogenetic studies indicate that Type I and Type II proteins have evolved differently (De Bodt *et al.* 2003; Nam *et al.* 2004; Veron *et al.* 2007). Type II regulatory networks have facilitated the diversification of this group of MADS-box genes. Functional and interaction analyses of Type I proteins should reveal the role of regulatory networks in the evolution of Type I proteins and the morphological innovations they underlie. A majority of the MADS-box genes that have been functionally characterized are Type II MIKCc. Surprisingly not all of the Type II MIKCc genes have been functionally characterized even in Arabidopsis. Recent functional analyses of an Arabidopsis B-sister MADS-box gene, *GORDITA*, has shown a novel role for a MADS-box gene in the regulation of organ growth (Prasad *et al.* 2010). There have been many Type II MIKCc studies performed across the land plants however, little is known about the Type II MIKC* and Type I MADS-box genes. To understand the role of MADS-box genes in the land plants, genetic functional systems need to be developed not only in additional angiosperm species but also in lycophyte, fern and gymnosperm species. Broad comparative functional analyses of Type II MIKC* and Type I MADS-box proteins across the land plants will help to elucidate their role in the evolution and development of land

plants.

REFERENCES

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

Adamczyk BJ, Fernandez DE (2009) MIKC* MADS domain heterodimers are required for pollen maturation and tube growth in Arabidopsis. *Plant Physiology* **149**, 1713-1723

AGI-Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815

Alvarez-Buylla ER, Pelaz S, Liljegren SJ, Gold SE, Burgeff C, Ditta GS, Ribas de Pouplana L, Martinez-Castilla L, Yanofsky MF (2000) An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proceedings of the National Academy of Sciences USA* **97**, 5328-5333

Ambrose BA, Lerner DR, Ciceri P, Padilla CM, Yanofsky MF, Schmidt RJ (2000) Molecular and genetic analyses of the silky1 gene reveal conservation in floral organ specification between eudicots and monocots. *Molecular Cell* **5**, 569-579

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

Becker A, Theissen G (2003) The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Molecular Phylogenetics and Evolution* **29**, 464-489

Bemer M, Wolters-Arts M, Grossniklaus U, Angenent GC (2008) The MADS domain protein DIANA acts together with AGAMOUS-LIKE80 to specify the central cell in Arabidopsis ovules. *The Plant Cell* **20**, 2088-2101

Bowman JL (1997) Evolutionary conservation of angiosperm flower development at the molecular and genetic levels. *Journal of Biosciences* **22**, 515-527

Bowman JL, Smyth DR, Meyerowitz EM (1989) Genes directing flower development in Arabidopsis. *The Plant Cell* **1**, 37-52

Carlsbecker A, Sundstrom J, Tandre K, Englund M, Kvarnheden A, Johanson U, Engstrom P (2003) The *DAL10* gene from Norway spruce (*Picea abies*) belongs to a potentially gymnosperm-specific subclass of MADS-box genes and is specifically active in seed cones and pollen cones. *Evolution and Development* **5**, 551-561

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

Davidson EH, Rast JP, Oliveri P, Ransick A, Calestani C, Yuh CH, Minokawa T, Amore G, Hinman V, Arenas-Mena C, Otim O, Brown CT, Livi CB, Lee PY, Revilla R, Rust AG, Pan Z, Schilstra MJ, Clarke PJ, Arnone MI, Rowen L, Cameron RA, McClay DR, Hood L, Bolouri H (2002) A genomic regulatory network for development. *Science* **295**, 1669-1678

Davies B, Schwarz-Sommer Z (1994) Control of floral organ identity by homeotic MADS-box transcription factors. *Results and Problems in Cell Differentiation* **20**, 235-258

Day RC, Herridge RP, Ambrose BA, Macknight RC (2008) Transcriptome analysis of proliferating Arabidopsis endosperm reveals biological implications for the control of syncytial division, cytokinin signaling, and gene expression regulation. *Plant Physiology* **148**, 1964-1984

De Bodt S, Raes J, Florquin K, Rombauts S, Rouze P, Theissen G, Van de Peer Y (2003) Genomewide structural annotation and evolutionary analysis of the type I MADS-box genes in plants. *Journal of Molecular Evolution* **56**, 573-586

de Folter S, Immink RG, Kieffer M, Parenicova L, Henz SR, Weigel D, Busscher M, Kooiker M, Colombo L, Kater MM, Davies B, Angenent GC (2005) Comprehensive interaction map of the Arabidopsis MADS Box transcription factors. *The Plant Cell* **17**, 1424-1433

Egea-Cortines M, Saedler H, Sommer H (1999) Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in *Antirrhinum majus*. *The EMBO Journal* **18**, 5370-5379

Espinosa-Soto C, Padilla-Longoria P, Alvarez-Buylla ER (2004) A gene regulatory network model for cell-fate determination during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles. *The Plant Cell* **16**, 2923-2939

Ferrandiz C, Gu Q, Martienssen R, Yanofsky MF (2000) Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. *Development* **127**, 725-734

Friedman WE, Williams JH (2003) Modularity of the angiosperm female gametophyte and its bearing on the early evolution of endosperm in flower-

ing plants. *Evolution* **57**, 216-230

Galant R, Carroll SB (2002) Evolution of a transcriptional repression domain in an insect Hox protein. *Nature* **415**, 910-913

Gehring WJ (2005) New perspectives on eye development and the evolution of eyes and photoreceptors. *Journal of Heredity* **96**, 171-184

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R, Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296**, 92-100

Gregis V, Sessa A, Colombo L, Kater MM (2006) AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis. *The Plant Cell* **18**, 1373-1382

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

Gu Q, Ferrandiz C, Yanofsky MF, Martienssen R (1998) The FRUITFULL MADS-box gene mediates cell differentiation during Arabidopsis fruit development. *Development* **125**, 1509-1517

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

Hasebe M, Wen CK, Kato M, Banks JA (1998) Characterization of MADS homeotic genes in the fern *Ceratopteris richardii*. *Proceedings of the National Academy of Sciences USA* **95**, 6222-6227

Heck GR, Perry SE, Nichols KW, Fernandez DE (1995) AGL15, a MADS domain protein expressed in developing embryos. *The Plant Cell* **7**, 1271-1282

Henschel K, Kofuji R, Hasebe M, Saedler H, Munster T, Theissen G (2002) Two ancient classes of MIKC-type MADS-box genes are present in the moss *Physcomitrella patens*. *Molecular Biology and Evolution* **19**, 801-814

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

Honys D, Twell D (2004) Transcriptome analysis of haploid male gametophyte development in Arabidopsis. *Genome Biology* **5**, R85

Irish VF, Litt A (2005) Flower development and evolution: gene duplication, diversification and redeployment. *Current Opinion in Genetics and Development* **15**, 454-460

Jager M, Hassanin A, Manuel M, Le Guyader H, Deutsch J (2003) MADS-box genes in *Ginkgo biloba* and the evolution of the AGAMOUS family. *Molecular Biology and Evolution* **20**, 842-854

Kang IH, Steffen JG, Portereiko MF, Lloyd A, Drews GN (2008) The AGL62 MADS domain protein regulates cellularization during endosperm development in Arabidopsis. *The Plant Cell* **20**, 635-647

Karol KG, McCourt RM, Cimino MT, Delwiche CF (2001) The closest living relatives of land plants. *Science* **294**, 2351-2353

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

Kempin SA, Savidge B, Yanofsky MF (1995) Molecular basis of the cauliflower phenotype in Arabidopsis. *Science* **267**, 522-525

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

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

Kramer EM, Jaramillo MA, Di Stilio VS (2004) Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS box genes in angiosperms. *Genetics* **166**, 1011-1023

Krizek BA, Fletcher JC (2005) Molecular mechanisms of flower development: an armchair guide. *Nature Review Genetics* **6**, 688-698

Krogan NT, Ashton NW (2000) Ancestry of Plant MADS-box genes revealed by bryophyte (*Physcomitrella patens*) homologues. *New Phytologist* **147**, 505-517

Leseberg CH, Li A, Kang H, Duvall M, Mao L (2006) Genome-wide analysis of the MADS-box gene family in *Populus trichocarpa*. *Gene* **378**, 84-94

Leseberg CH, Eissler CL, Wang X, Johns MA, Duvall MR, Mao L (2008) Interaction study of MADS-domain proteins in tomato. *Journal of Experimental Botany* **59**, 2253-2265

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

Litt A, Irish VF (2003) Duplication and diversification in the *APETALA1/FRUITFULL* floral homeotic gene lineage: implications for the evolution of

floral development. *Genetics* **165**, 821-833

Ma H, Yanofsky MF, Meyerowitz EM (1991) AGL1-AGL6, an Arabidopsis gene family with similarity to floral homeotic and transcription factor genes. *Genes and Development* **5**, 484-495

Menand B, Calder G, Dolan L (2007) Both chloronemal and caulonemal cells expand by tip growth in the moss *Physcomitrella patens*. *Journal of Experimental Botany* **58**, 1843-1849

Mendoza L, Thieffry D, Alvarez-Buylla ER (1999) Genetic control of flower morphogenesis in *Arabidopsis thaliana*: a logical analysis. *Bioinformatics* **15**, 593-606

Meyerowitz EM, Bowman JL, Brockman LL, Drews GN, Jack T, Sieburth LE, Weigel D (1991) A genetic and molecular model for flower development in *Arabidopsis thaliana*. *Development* **1 (Suppl)**, 157-167

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

Munster T, Pahnke J, Di Rosa A, Kim JT, Martin W, Saedler H, Theissen G (1997) Floral homeotic genes were recruited from homologous MADS-box genes preexisting in the common ancestor of ferns and seed plants. *Proceedings of the National Academy of Sciences USA* **94**, 2415-2420

Nam J, dePamphilis CW Ma, H, Nei M (2003) Antiquity and evolution of the MADS-box gene family controlling flower development in plants. *Molecular Biology and Evolution* **20**, 1435-1447

Nam J, Kim J, Lee S, An G, Ma H, Nei M (2004) Type I MADS-box genes have experienced faster birth-and-death evolution than type II MADS-box genes in angiosperms. *Proceedings of the National Academy of Sciences USA* **101**, 1910-1915

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

Olson EN (2006) Gene regulatory networks in the evolution and development of the heart. *Science* **313**, 1922-1927

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

Passmore S, Maine GT, Elble R, Christ C, Tye BK (1988) *Saccharomyces cerevisiae* protein involved in plasmid maintenance is necessary for mating of MAT alpha cells. *Journal of Molecular Biology* **204**, 593-606

Pelaz S, Ditta GS, Baumann E, Wisman E, Yanofsky MF (2000) B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature* **405**, 200-203

Portereiko MF, Lloyd A, Steffen JG, Punwani JA, Otsuga D, Drews GN (2006) AGL80 is required for central cell and endosperm development in Arabidopsis. *The Plant Cell* **18**, 1862-1872

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

Purugganan MD, Rounsley SD, Schmidt RJ, Yanofsky MF (1995) Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. *Genetics* **140**, 345-356

Rensing SA, Lang D, Zimmer AD, Terry A, Salamov A, Shapiro H, Nishiyama T, Perroud PF, Lindquist EA, Kamisugi Y, Tanahashi T, Sakakibara K, Fujita T, Oishi K, Shin IT, Kuroki Y, Toyoda A, Suzuki Y, Hashimoto S, Yamaguchi M, Sugano S, Kohara Y, Fujiyama A, Anterola A, Aoki S, Ashton N, Barbazuk WB, Barker E, Bennetzen JL, Blankenship R, Cho SH, Dutcher SK, Estelle M, Fawcett JA, Gundlach H, Hanada K, Heyl A, Hicks KA, Hughes J, Lohr M, Mayer K, Melkozernov A, Murata T, Nelson DR, Pils B, Prigge M, Reiss B, Renner T, Rombauts S, Rushton PJ, Sanderfoot A, Schween G, Shiu SH, Stueber K, Theodoulou FL, Tu H, Van de Peer Y, Verrier PJ, Waters E, Wood A, Yang L, Cove D, Cuming AC, Hasebe M, Lucas S, Mishler BD, Reski R, Grigoriev IV, Quatrano RS, Boore JL (2008) The Physcomitrella genome reveals evolutionary insights into the conquest of land by plants. *Science* **319**, 64-69

Riano-Pachon DM, Correa LG, Trejos-Espinosa R, Mueller-Roeber B (2008) Green transcription factors: a chlamydomonas overview. *Genetics* **179**, 31-39

Riechmann JL, Meyerowitz EM (1997) Determination of floral organ identity by Arabidopsis MADS domain homeotic proteins AP1, AP3, PI, and AG is independent of their DNA-binding specificity. *Molecular Biology of the Cell* **8**, 1243-1259

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

Riese M, Faigl W, Quodt V, Verelst W, Matthes A, Saedler H, Munster T (2005) Isolation and characterization of new MIKC*-Type MADS-box genes from the moss *Physcomitrella patens*. *Plant Biology* (Stuttgart) **7**, 307-314

Rijpkema AS, Zethof J, Gerats T, Vandenbussche M (2009) The petunia *AGL6* gene has a *SEPALLATA*-like function in floral patterning. *The Plant Journal* **60**, 1-9

Rounsley SD, Ditta GS, Yanofsky MF (1995) Diverse roles for MADS box genes in Arabidopsis development. *The Plant Cell* **7**, 1259-1269

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

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

Shore P, Sharrocks AD (1995) The MADS-box family of transcription factors. *European Journal of Biochemistry* **229**, 1-13

Sommer H, Beltran JP, Huijser P, Pape H, Lonnig WE, Saedler H, Schwarz-Sommer Z (1990) *Deficiens*, a homeotic gene involved in the control of flower morphogenesis in *Antirrhinum majus*: the protein shows homology to transcription factors. *The EMBO Journal* **9**, 605-613

Steffen JG, Kang IH, Portereiko MF, Lloyd A, Drews GN (2008) AGL61 interacts with AGL80 and is required for central cell development in Arabidopsis. *Plant Physiology* **148**, 259-268

Svensson ME, Johannesson H, Engstrom P (2000) The LAMB1 gene from the clubmoss, *Lycopodium annotinum*, is a divergent MADS-box gene, expressed specifically in sporogenic structures. *Gene* **253**, 31-43

Swalla BJ (2006) Building divergent body plans with similar genetic pathways. *Heredity* **97**, 235-243

Tanabe Y, Uchida M, Hasebe M, Ito M (2003) Characterization of the *Selaginella remotifolia* MADS-box gene. *Journal of Plant Research* **116**, 71-75

Tanabe Y, Hasebe M, Sekimoto H, Nishiyama T, Kitani M, Henschel K, Munster T, Theissen G, Nozaki H, Ito M (2005) Characterization of MADS-box genes in charophycean green algae and its implication for the evolution of MADS-box genes. *Proceedings of the National Academy of Sciences USA* **102**, 2436-2441

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

Theissen G, Becker A, Di Rosa A, Kanno A, Kim JT, Munster T, Winter KU, Saedler H (2000) A short history of MADS-box genes in plants. *Plant Molecular Biology* **42**, 115-149

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, Putnam N, Ralph S, Rombauts S, Salamov A, Schei J, Sterck, L, Aerts A, Bhalerao RR, Bhalerao RP, Blaudez D, Boerjan W, Brun A, Brunner A, Busov V, Campbell M, Carlson J, Chalot M, Chapman J, Chen GL, Cooper D, Coutinho PM, Couturier J, Covert S, Cronk Q, Cunningham R, Davis J, Degroeve S, Dejardin A, Depamphilis C, Detter J, Dirks B, Dubchak I, Duplessis S, Ehlting J, Ellis B, Gendler K, Goodstein D, Gribskov M, Grimwood J, Groover A, Gunter L, Hamberger B, Heinze B, Helariutta Y, Henrissat B, Holligan D, Holt R, Huang W, Islam-Faridi N, Jones S, Jones-Rhoades M, Jorgensen R, Joshi C, Kangasjarvi J, Karlsson J, Kelleher C, Kirkpatrick R, Kirst M, Kohler A, Kalluri U, Larimer F, Leebens-Mack J, Leple JC, Locascio P, Lou Y, Lucas S, Martin F, Montanini B, Napoli C, Nelson DR, Nelson C, Nieminen K, Nilsson O, Pereda V, Peter G, Philippe R, Pilate G, Poliakov A, Razumovskaya J, Richardson P, Rinaldi C, Ritland K, Rouze P, Ryaboy D, Schmutz J, Schrader J, Segerman B, Shin H, Siddiqui,A, Sterky F, Terry A, Tsai CJ, Uberbacher E, Unneberg P, Vahala J, Wall K, Wessler S, Yang G, Yin T, Douglas C, Marra M, Sandberg G, Van de Peer Y, Rokhsar D (2006) The genome of black cottonwood, Populus trichocarpa (Torr. & Gray). *Science* **313**, 1596-1604

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

Veron AS, Kaufmann K, Bornberg-Bauer E (2007) Evidence of interaction network evolution by whole-genome duplications: a case study in MADS-box proteins. *Molecular Biology and Evolution* **24**, 670-678

Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W, Giovannoni J (2002) A MADS-box gene necessary for fruit ripening at the tomato ripening-inhibitor (rin) locus. *Science* **296**, 343-346

Walia H, Josefsson C, Dilkes B, Kirkbride R, Harada J, Comai L (2009) Dosage-dependent deregulation of an AGAMOUS-LIKE gene cluster contributes to interspecific incompatibility. *Current Biology* **19**, 1128-1132

Winter KU, Becker A, Munster T, Kim JT, Saedler H, Theissen G (1999) MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. *Proceedings of the National Academy of Sciences USA* **96**, 7342-7347

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

Yoo SK, Lee JS, Ahn JH (2006) Overexpression of AGAMOUS-LIKE 28 (AGL28) promotes flowering by upregulating expression of floral promoters within the autonomous pathway. *Biochemical and Biophysical Research Communications* **348**, 929-936

Yu J, Hu S, Wang J, Wong GK, Li S, Liu B, Deng Y, Dai L, Zhou Y, Zhang X, Cao M, Liu J, Sun J, Tang J, Chen Y, Huang X, Lin W, Ye C, Tong W, Cong L, Geng J, Han Y, Li L, Li W, Hu G, Li J, Liu Z, Qi Q, Li T, Wang X, Lu H, Wu T, Zhu M, Ni P, Han H, Dong W, Ren X, Feng X, Cui P, Li X, Wang H, Xu X, Zhai W, Xu Z, Zhang J, He S, Xu J, Zhang K, Zheng X, Dong J, Zeng W, Tao L, Ye J, Tan J, Chen X, He J, Liu D, Tian W, Tian C, Xia H, Bao Q, Li G, Gao H, Cao T, Zhao W, Li P, Chen W, Zhang Y, Hu J, Liu S, Yang J, Zhang G, Xiong Y, Li Z, Mao L, Zhou C, Zhu Z, Chen R, Hao B, Zheng W, Chen S, Guo W, Tao M, Zhu L, Yuan L, Yang H (2002) A draft sequence of the rice genome (Oryza sativa L. ssp. indica). *Science* **296**, 79-92

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