INVITED REVIEW

Molecular Mechanisms Underlying Origin and Diversification of the Angiosperm Flower

GUENTER THEISSEN* and RAINER MELZER
Friedrich-Schiller-Universitaet Jena, Lehrstuhl fuer Genetik, Philosophenweg 12, D-07743 Jena, Germany

Received: 2 March 2007 Returned for revision: 25 April 2007 Accepted: 31 May 2007 Published electronically: 31 July 2007

INTRODUCTION

Evolution of the angiosperm flower: news and views from Darwin until now

Imagine roses, tulips, lilies and orchids – we all feel quite familiar with flowers, don’t we? Many will be quite surprised, therefore, to learn that the modes and mechanisms of the origin and early diversification of the flowering plants (angiosperms) is still one of the most hotly debated matters of evolutionary biology.

The angiosperms provide us, directly or indirectly, with the vast majority of human food (such as vegetables and fruits) as well as with a lot of other economically important products (e.g. use in the preparation of clothes, furniture and drugs), and they also have great aesthetic value as ornamental plants. In addition, the contributions of angiosperms to biodiversity in terrestrial ecosystems can hardly be overestimated. No wonder, therefore, that botanists and evolutionary biologists made considerable efforts to clarify the origin and diversification of the flowering plants. The apparently ‘sudden’ origin and rapid early morphological radiation of the angiosperms, as revealed by the fossil record, already intrigued Charles Darwin, who considered it a ‘perplexing phenomenon’ and even an ‘abominable mystery’ (Crepet, 1998, 2000).

Like most other real mysteries, explaining the early evolution of angiosperms proved difficult, despite considerable efforts during the last one and a half centuries (Frohlich, 2003, 2006). Major reasons for this are a quite uninformative (and probably extremely incomplete) fossil record and the great morphological gap between the flowers of angiosperms and the reproductive structures of gymnosperms, angiosperms’ closest relatives; gymnosperms comprise extant conifers (including well-known taxa such as spruce and pine species), gnetophytes, cycads, Ginkgo, and diverse extinct groups including Bennettitales, Cordaitales, Cystospermuses and Glossopterides (Doyle, 1998). The morphological gap between gymnosperms and angiosperms leads to problems with homology assignments between reproductive organs from flowering plants and their putative ancestors and thus hampers an understanding of the transition from gymnosperm reproductive cones to angiosperm flowers (Frohlich, 2003).

Flowers are characteristic features of angiosperms, and their origin is a major aspect of the ‘abominable mystery’. Given the importance of flowers one wonders that botanists have difficulties to define both precisely and comprehensively what actually a flower is. Bateman et al. (2006) have listed about a dozen of historical circumscriptions, and these authors as well as Baum and Hileman (2006) came up with updated definitions of a flower. As a consensus of these contemporary attempts, one may define a flower as a determinate, compressed, bisexual reproductive axis composed of megasporangia (carpels), microsporangia (stamens) and a sterile perianth composed of at least one sterile laminar organ. This is obviously not a comprehensive

* For correspondence. E-mail guenter.theissen@uni-jena.de
definition, since there are numerous types of flowers that lack one or more of these organ types, probably due to secondary simplification (e.g. unisexual flowers with only stamens or carpels).

This leads to the question as to which of these features represent the key characteristics of angiosperm flowers (for a detailed discussion of the synapomorphies of the angiosperms, see Donoghue and Doyle, 1989). The presence of carpels enclosing the ovules is generally considered an essential character that distinguishes typical angiosperm flowers from gymnosperms’ reproductive cones (Crane et al., 1995; Endress, 2001; Stuessy, 2004). Another key feature of angiosperm flowers is a trait termed hermaphroditism, or bisexuality, describing the fact that male and female reproductive organs are united in one structure (or secondarily separated, as in the unisexual flowers of monoecious and dioecious angiosperms), while they might be primarily separated in different structures in gymnosperms. Moreover, flowers have a compressed reproductive axis, while gymnosperms have elongate ones (Baum and Hileman, 2006). The presence of a perianth surrounding the reproductive organs, often including attractive organs of petaloid appearance (petals or tepals) is considered yet another typical feature distinguishing angiosperm flowers from gymnosperm cones. Often double fertilization yielding a nutritive endosperm is also considered an important angiosperm feature (Stuessy, 2004; Kramer and Jaramillo, 2005). Obviously, a flower is not a simple characteristic, but a complex of innovations (Baum and Hileman, 2006).

Almost all terrestrial environments are currently dominated by flowering plants. To what extent the evolution of the flower contributed to the evolutionary ‘success’ of the angiosperms is unknown, but there is good reason to assume that it was of quite some importance. In addition to allowing efficient pollination by animals, thereby permitting population densities lower than in gymnosperms without inbreeding depression and extinction (Regal, 1977), flowers also ‘catalyse’ seed dispersal via the formation of fruits, also often employing animal vectors. Flowers also increase the potential for reproductive isolation, e.g. by changes in pollinator specificity or flowering time (Grant, 1971).

The tempo and mode of morphological changes by which the angiosperm flower originated has remained enigmatic. Based on circumstantial evidence, several authors have argued that bisexuality may have been one of the first steps during flower origin (see below, and Theissen et al., 2002; Theißen and Becker, 2004; Baum and Hileman, 2006). Due to a lack of informative fossils, the other steps are even more speculative. Stuessy (2004) hypothesized that angiosperms may have evolved slowly from seed ferns in the Jurassic, beginning first with the carpel, followed later by double fertilization, and lastly by the appearance of flowers. This series of events may have well taken more than 100 million years to complete, before an ‘explosive’ evolutionary diversification may have set in when the final combination of the essential angiosperm features was achieved (Stuessy, 2004). Ignoring the question of carpel origin, Baum and Hileman (2006) proposed that evolution of a bisexual axis via a gynonomocious intermediate was the first step during flower origin, followed by evolution of floral axis compression and determinacy as step two, evolution of a petaloid perianth by sterilization of the outer stamens in step three, and origin of the dimorphic perianth of core eudicots in step four. What makes this scenario especially intriguing is the fact that for each of these steps an underlying developmental genetic mechanism is suggested by Baum and Hileman (2006). Thus the scenario outlined by these authors can in principle be experimentally tested by comparative developmental genetic studies in gymnosperms and angiosperms. This, however, is much easier said than done, given the poor performance of extant gymnosperms as developmental genetic model systems; they are all woody species requiring lots of growth space and many years of vegetative development until they reach the reproductive stage.

Various aspects of the evolutionary origin of the angiosperm flower have been reviewed by several authors during recent years (Crepet, 1998, 2000; Frohlich, 1999, 2003, 2006; Frohlich and Parker, 2000; Ma and dePamphilis, 2000; Albert et al., 2002; Theissen et al., 2002; Stuessy, 2004; Theißen and Becker, 2004; Theissen, 2005a; Theißen and Kaufmann, 2006; Baum and Hileman, 2006; Bateman et al., 2006; Doyle, 2006). With the accumulating knowledge about the molecular genetic basis of flower development in some model plants the proximate causes and molecular mechanisms of floral evolution have found increasing interest. Here we focus on an especially well-understood aspect of flowers from a developmental genetic point of view: organ identity. We outline how two leading models explaining the specification of floral organ identity during individual plant development are helping to better understand flower origin and diversification.

Evolution of seed plants and the ancestral flower of crown group angiosperms

Understanding the phylogeny of seed plants (spermatophytes), a clade comprising angiosperms + gymnosperms, is an important prerequisite for understanding the evolution of the reproductive structures of these taxa, including flowers. Much has been learned about the phylogeny of seed plants in recent years by the use of molecular markers, even though some critical aspects have remained unresolved and controversial (Bateman et al., 2006; Frohlich, 2006).

Extant gymnosperms are morphologically very diverse and hence have usually been considered being paraphyletic, with gnetophytes (Gnetum, Ephedra and Welwitschia) often regarded as the sister group of the angiosperms (Doyle, 1998). It thus came as a surprise that the vast majority of studies using molecular markers found moderate to strong support for extant gymnosperms being monophyletic (see, for example, Chaw et al., 1997, 2000; Soltis et al. 1999a; Winter et al., 1999; Bowe et al., 2000; Frohlich and Parker, 2000; reviewed by Frohlich, 2006). Extant gymnosperms and angiosperms may thus be sister groups that, as also suggested by molecular evidence (see e.g. Goremykin et al., 1997), separated roughly about
300 million years ago. These hypotheses, however, are not easy to reconcile with the fact that the most ancient reliable angiosperm fossils, i.e. those of *Archaefructus*, are just about 130 million years old (Sun et al., 2002; Friis et al., 2003), because this means there would have been about 170 million years of evolution during which the lineage that led to extant angiosperms left no recognizable traces in the fossil record (Bateman et al., 2006). Thus especially some (but not all) paleobotanists do not accept the molecular-derived hypotheses about seed plant phylogeny, and consider the issue as unresolved (for an overview, see Bateman et al., 2006; Frohlich, 2006).

In any case, the inability to unambiguously identify an (extinct or extant) group of gymnosperms as the closest relatives of the angiosperms implies that the gap between the two kinds of seed plants could not be bridged yet. The deep phylogeny of extant angiosperms itself is also still quite controversial [compare, for example, Soltis and Soltis (2004) and Soltis et al. (2004) with Goremykin et al. (2003, 2004)], but it seems that we are much closer to a reasonable consensus [especially in case of deep seed plant phylogeny (Goremykin and Hellwig, 2006; Bateman et al., 2006).

There is evidence that either *Amborella trichopoda*, or a clade of *Amborella* together with Nymphaeales (water lilies) represent the most basal angiosperms (see, for example, Qiu et al., 1999; Soltis et al., 1999b; Barkman et al., 2000; Graham and Olmstead, 2000; for a review see Kuzoff and Gasser, 2000; Soltis and Soltis, 2003; Soltis et al. 2005). As the next branch, this clade or grade is then followed by Austrobaileyales, a monophyletic group unifying Schisandraceae (now also including Illiciaceae; Frohlich, 2006), Trimeniaceae and Austrobaileyales. [The grade of probably most basal angiosperms was originally termed ‘ANITA’ (Qiu et al., 1999), and has recently been named ‘ANA’ (*Amborella*, Nymphaeales, Austrobaileyales) (Frohlich, 2006).]

Assuming that *Amborella*, Nymphaeales and Austrobaileyales indeed represent the most basal angiosperms, the flowers at the base of crown group angiosperms can be reconstructed by consideration of floral structures within this grade. These ‘ancestral flowers’ (Frohlich, 2006) were probably already hermaphroditic and had an undifferentiated perianth, in which organs were arranged in more than two cycles or a spiral. Given their completeness and complexity these flowers at the base of crown group angiosperms, hence probably represent already a quite late, rather than initial stage during the origin of the flower as we know it. However, flowers with differentiated sepals and petals probably evolved even later during the evolution of angiosperms (Albert et al., 1998; Kuzoff and Gasser, 2000; Ronse De Craene et al., 2003; Soltis et al., 2005).

In the framework of our current view on angiosperm phylogeny, the clade comprising all flowering plants above the basally diverging lineages is called the ‘euangiosperms’, including the magnoloids, monocots, Chloranthaceae and eudicots. The phylogenetic relationships of these groups are unresolved. The eudicots comprise a grade of successive branches, with Ranunculales, including Papaveraceae, as sister to all other eudicots, and a large clade of ‘core eudicots’ containing the majority of all angiosperm species, including Caryophyllales (to which Polygonaceae belong), Rosids (to which Brassicaceae belong), Asterids, and several additional lineages (Fig. 1). The major eudicot model plant thale cress (*Arabidopsis thaliana*; henceforth termed Arabidopsis) belongs to the plant family Brassicaceae within the eurosids (APG II, 2003; Soltis and Soltis, 2003).

From the ABC model to the quartet model of floral organ identity

Eudicots represent the largest group of flowering plants, comprising about 75% of angiosperm species (Buzgo et al., 2005). Eudicots have relatively standardized flowers that typically consist of four different classes of organs arranged in four (or more) whorls at the tip of a floral shoot. The first, outermost whorl usually consists of green sepals resembling vegetative leaves. The second whorl is composed of often relatively large, coloured and showy petals. The third whorl contains the stamens, i.e. the male reproductive organs which produce the pollen. Finally, the fourth, innermost whorl contains the carpels, i.e. the female reproductive organs, which are often fused and inside of which, after fertilization, the ovules develop into seeds.

The structures of mature sepals, petals, stamens and carpels differ dramatically. Nevertheless, each floral organ starts its development as a little bulge generated by anticlinal and periclinal divisions of undifferentiated cells of the floral meristem. Thus each cell in the developing floral organ primordium somehow ‘recognizes’ its position within the flower, and differentiates accordingly into a cell type that is appropriate for the specific organ.

Based on the study of homeotic mutants in which the identity of floral organs is changed, genetic models were developed that explain how the different floral organs acquire their specific ‘identities’ during development (reviewed by Theissen, 2001). In Arabidopsis, homeotic mutants come in three classes, A, B and C. Ideal class A mutants have carpels rather than sepals in the first whorl, and stamens instead of petals in the second whorl. Class B mutants have sepals rather than petals in the second whorl, and carpels replace stamens in the third whorl. Class C mutants develop petals rather than sepalas in the second whorl, and carpels replace stamens in the fourth whorl. In addition, the flowers of class C mutants grow indeterminately, i.e. there is continued production of mutant floral organs inside the 4th whorl; this is in contrast to wild-type flowers, whose development ceases after carpel formation.

The phenotypes of class A, B and C mutants proved extremely informative; they suggested that flower development is sculpted by homeotic selector genes (also termed ‘floral organ identity genes’). These genes act as major developmental switches that control the entire genetic programme required for the development of a particular organ. The activities and interactions of floral homeotic genes have been described in several quite similar models
Haughn and Somerville, 1988; Schwarz-Sommer et al., 1990), of which one, the ‘ABC model’ (Coen and Meyerowitz, 1991) became widely known (Fig. 1). This ‘classical ABC model’ maintains that organ identity in each floral whorl is determined by a unique combination of three organ identity gene activities, called A, B and C. Expression of class A genes alone specifies the formation of sepals. The combination AB specifies the development of petals, and the combination BC specifies stamen formation. Expression of C alone determines the development of carpels. In order to explain the three classes of floral homeotic mutants, the ABC model proposes that the class A and class C genes negatively regulate each other, so that the class C gene activity becomes expressed throughout the flower when the class A gene function is compromised, and vice versa (for reviews of the ABC model, see Weigel and Meyerowitz, 1994; Theißen, 2001).

In Arabidopsis, class A genes are represented by APETALA1 (AP1) and APETALA2 (AP2), class B genes by APETALA3 (AP3) and PISTILLATA (PI), and class C genes by AGAMOUS (AG). Molecular cloning of these genes revealed that, except AP2, they all represent MIKC-type MADS-box genes encoding transcription factors (MADS-domain proteins) (for a review see Theißen, 2001). Thus the products of the ABC genes probably control the transcription of other genes (‘target genes’).
whose products are directly or indirectly involved in the formation or function of floral organs.

MIKC-type proteins have a characteristic domain structure, including, from N- to C-terminus, a MADS (M-), intervening (I-), keratin-like (K-) and C-terminal (C-) domain (Münter et al., 1997). In all kinds of MADS-domain proteins, the MADS-domain is by far the most highly conserved region; it is the major determinant of DNA binding, but it also performs functions in dimerization of MADS-domain proteins and in the binding of accessory factors. MADS-domain proteins bind to DNA sites with similarity to the consensus sequence 5’-CC(A/T)6GG-3’, termed a ‘CARG-box’ (for ‘CC-A rich-GG’). The I-domain, located directly downstream of the MADS-domain, is relatively variable in length and not very strongly conserved. At least in some MIKC-type proteins, it constitutes an important determinant for the selective formation of DNA-binding dimers. The K-domain is characterized by a conserved, quite regular spacing of hydrophobic residues, which is proposed to allow for the formation of amphiphatic helices. These are assumed to interact with amphiphatic helices of other K-domain-containing proteins via the formation of coiled-coils to promote protein dimerization or multimeric complex formation (Kaufmann et al., 2005). The most variable region is the C-domain at the C-terminal end of MIKC-type proteins. In some MIKC-type proteins it is involved in transcriptional activation, or in the formation of multimeric complexes (structural and phylogenetic aspects of MIKC-type proteins have been reviewed in detail by Kaufmann et al., 2005).

Despite its heuristic value, the ABC model has two major shortcomings: mutant and transgenic studies revealed that the ABC genes are required, but not sufficient for the specification of floral organ identity. Moreover, the ABC model does not provide a molecular mechanism for the specification of organ identity by the interaction of floral homeotic genes. Development of the ‘ABCD’ and the ‘floral quartet’ models allowed both of these shortcomings to be overcome. Based on the analysis of multiple mutants generated via ‘reverse genetics’, the ABC model was first extended to an ‘ABCD model’ by addition of class D genes that are closely related to class C genes and specify ovule identity (Angenent and Colombo, 1996). Further consideration of class D genes is not required in the framework of this review.

Knowing about another class of closely related MIKC-type MADS-box genes – originally known as AGL2-like genes (Theißen et al., 1996), but later renamed into SEPALLATA (SEP)-like genes (Pelaz et al., 2000), is essential here, however. Arabidopsis has four different SEP-like genes, termed SEP1–SEP4. While all single and the double mutants of SEP revealed only very weak, if any, deviations from wild-type phenotype, in sep1 sep2 sep3 triple mutants the organs in all whorls of the flower develop into sepals, and flower development becomes indeterminate (Pelaz et al., 2000); in sep1 sep2 sep3 sep4 quadruple mutants vegetative leaves rather than sepals develop in all whorls of indeterminate flowers (Ditta et al., 2004).

The SEP genes are still expressed in loss-of-function mutants of class B and class C genes, and the initial expression patterns of B and C class genes are not altered in the sep1 sep2 sep3 triple mutant (Pelaz et al., 2000). Therefore, the SEP genes are considered as yet another class of floral organ identity genes, termed class E genes (Theißen, 2001), required for the development of all categories of floral organs.

According to the ABCDE model class A + E genes are required to specify sepals, A + B + E petals, B + C + E stamens, C + E carpels and D + E ovules (Theißen, 2001; Ditta et al., 2004). But by which molecular mechanism do the different floral homeotic genes interact?

All MADS-domain proteins tested bind to DNA in a sequence-specific way as dimers. Explaining the interaction of floral homeotic genes by dimerization of their gene products soon failed, however, because the expected dimers were not observed in respective studies (Riechmann et al., 1996; Fan et al., 1997). For example, the class B proteins AP3 and PI bind to DNA only as obligate AP3–PI heterodimers, but DNA-binding dimers between AP3 or PI and AP1 (class A) or AG (class C) proteins were not observed (Riechmann et al., 1996). An obvious way to explain the combinatorial interactions would be that class A and class B proteins (for specifying petals) or class B and class C proteins (for specifying stamens) bind separately to different cis-regulatory elements in the promoters of the same target genes that are activated or repressed during petal or stamen formation, respectively. Alternatively, class A and class B (class B and class C) proteins could have distinct sets of target genes involved in petal (stamen) formation, so that the combinatorial interaction of the homeotic genes would be mechanistically realized only at the level of target genes, or even further downstream in the gene cascades involved.

Things turned out to be more interesting. An important clue was provided when Egea-Cortines et al. (1999) reported that DEFICIENS (DEF), GLOBOSA (GLO) and SQUAMOSA (SQUA) from snapdragon (Antirrhinum majus) – putative orthologues of AP3, PI and AP1, respectively – form multimeric complexes in electrophoretic mobility shift assays and yeast three-hybrid analyses. The authors hypothesized that the protein complex is actually a protein tetramer, composed of a DEF–GLO heterodimer and a SQUA–SQUA homodimer, in which the DEF–GLO and SQUA–SQUA dimers recognize different CARG-boxes (Egea-Cortines et al., 1999).

When Pelaz et al. (2000) reported that not only the ABC but also the SEP genes are required for the formation of petals, stamens and carpels, the available data about the mechanisms of floral organ specification were pulled together in the ‘floral quartet model’ (Theißen, 2001). It suggests that complexes of four floral homeotic proteins including SEP proteins control floral organ identity. According to the original model there is at least one unique quaternary complex for each type of floral organ (Theißen, 2001). These quaternary protein complexes might function as transcription factors by binding to CARG-boxes in the promoters of target genes, which they either activate or repress as appropriate for the development of the identities of the different floral organs (Theißen, 2001). There is evidence that in these complexes, the API or SEP proteins
provide the transcription-activation domain, while the other proteins might be important for organ-specificity of gene regulation (Honma and Goto, 2001). The floral quartet model suggests that two protein dimers of each tetramer recognize two different CARG-boxes, which might be brought into close vicinity by bending the DNA between the CARG-boxes (Egea-Cortines et al., 1999; Theißen, 2001; Theißen and Saedler, 2001), so the quartet (or tetramer) is actually a dimer of dimers.

The floral quartet model was soon corroborated by Honma and Goto (2001) who demonstrated the formation of the complexes postulated for stamens and petals, namely AP3/PI/AG/SEP and AP3/PI/AP1 (or SEP), respectively, in yeast three-hybrid and four-hybrid assays. Moreover, it was shown that ectopic co-expression of class A/E + B genes AP1 + AP3 + PI or SEP3 + AP3 + P1 in transgenic Arabidopsis plants leads to a reprogramming of supposed-to-be leaf primordia in a way that they develop into petaloid organs. AP1 and SEP3 can substitute each others which probably reflects their close evolutionary relationship (Becker and Theißen, 2003) and hence partial redundancy (Honma and Goto, 2001; Pelaz et al., 2001). The co-expression of the class B + C + E genes AP3 + P1 + AG + SEP3 converts vegetative leaves into stamens. Nonetheless, these data add not only support for the floral quartet model but also demonstrate that just a few proteins are both necessary and sufficient to superimpose petal and stamen identity upon a vegetative developmental programme.

Clearly, the empirical basis of the floral quartet model is still quite weak, especially since there is no direct evidence so far that multimeric complexes of floral homeotic proteins really exist, loop DNA, and regulate target gene expression in the nuclei of plant cells. Moreover, a number of important questions remain unanswered by the floral quartet model, for example, how target gene specificity is achieved by floral homeotic proteins (Melzer et al., 2006). Limitations and possible experimental tests of the floral quartet model are discussed by Theißen and Melzer (2006).

Nevertheless, since the floral quartet model, in contrast to the ABCDE model, demonstrates how the combinatorial interaction of the floral homeotic genes works mechanistically, it has been widely accepted and became the standard model for work on floral homeotic genes (see, for example, Jack, 2001; Winter et al., 2002a; Ferrario et al., 2003; Lee et al., 2003; Ferrario et al., 2004; Shchennikova et al., 2004; Krizek and Fletcher, 2005; Robles and Pelaz, 2005). The floral quartet model may thus facilitate research on flower development and evolution in a similar way as the ABC model did 10 years before.

In the following, we outline the impact of the floral quartet model and the ABC model on understanding flower origin and diversification, respectively.

**FLORAL QUARTETS AND FLOWER ORIGIN**

*Molecular aspects of flower origin: a primer*

In the eudicot and monocot model plants, but by inference probably also in all other angiosperms, the identity of floral organs is totally dependent on the uncompromised activity of floral homeotic genes. Clarifying the origin of these genes and of the interactions in which they and their protein products are involved in may thus provide us with valuable insights into the origin of the flower (Theißen and Saedler, 1995; Theißen et al., 2000, 2002). Characterization of the MADS-box gene families in mosses and ferns suggested that, despite the presence of numerous MIKC-type genes, orthologues of floral homeotic genes are absent in non-seed plants (for a review, see Theißen et al., 2000). However, putative orthologues of class B and class C floral homeotic genes were isolated from different conifers and the gnetophyte *Gnetum gnemon*; orthologues of class C floral homeotic genes were also identified in *Cycas* and *Ginkgo* (Tandre et al., 1995, 1998; Rutledge et al., 1998; Mouradov et al., 1999; Sundström et al., 1999; Winter et al., 1999; Theißen et al., 2000; Fukui et al., 2001; Jager et al., 2003; Theißen and Becker, 2004; Zhang et al., 2004). (More precisely, gymnosperms contain putative precursors of both class C and D genes, which are here considered as class C genes to simplify things, even though class C/D genes would be a more precise term.) This suggests that class B and class C genes were established by gene duplications and subsequent divergence in the lineage that led to extant spermatophytes, after the lineage that led to extant ferns had already branched off, i.e. 300–400 million years ago; note, however, that some molecular clock estimates suggested a much earlier origin of these genes (Nam et al., 2003).

The expression patterns of gymnosperm class B and C gene orthologues resemble those of class B and class C genes in angiosperms, with class C genes being generally expressed in reproductive organs (both male and female ones), while expression of class B genes is focused on male reproductive organs (for a review, see Theißen et al., 2000; Theißen and Becker, 2004). Class B and class C genes from diverse gymnosperms could partially or fully substitute for their angiosperm orthologues in different kinds of complementation and ectopic expression experiments, thus revealing conservation of gene function during more than 300 million years of seed plant evolution (Sundström and Engström, 2002; Winter et al., 2002a; Zhang et al., 2004). Together with the expression data in gymnosperms these findings suggest that the ABC system specifying floral organ identity in flowering plants evolved from a BC precursor system that was already established in the most recent common ancestor of extant seed plants about 300 million years ago (Winter et al., 1999). We hypothesize that the BC system specified female reproductive organs by the expression of class C genes, and male reproductive organs by the expression of both class C and class B genes (Fig. 1) (Winter et al., 1999; Theißen et al., 2002; Theißen and Becker, 2004). Thus differential expression of class B genes specifying male organ identity may represent the primary sex-determination mechanism of all seed plants (Winter et al., 1999). If true, switching from male to female organ identity, or vice versa, could, in principle, simply result from changes in the spatial or temporal expression of one or more class B genes,
assumed that these genes encode transcription factors that control, directly or indirectly, all the target genes required to bring about male rather than female organ identity during development.

Based on such a simple switch mechanism, novel hypotheses on the origin of the bisexuality of flowers were developed; these hypotheses start with unisexual axes as existing in most extant gymnosperms (Theissen et al., 2002). The ‘out-of-male’ scenario maintains that bisexual flowers originated from a male gymnosperm cone, and that reduction of class B gene expression in the upper region of the male cone led to the development of female instead of male reproductive units in this upper region (Theissen et al., 2002). The ‘out-of-female’ scenario assumes that flowers originated from a female cone. In this scenario ectopic expression of class B genes in the basal region of the female cone led to the development of male rather than female reproductive units in this basal region (Theissen et al., 2002). Under both hypotheses a perianthless flower-like structure with male reproductive units in the basal region (outer whorls) and female reproductive units in the apical region (inner whorls) would have been established. Mutant analysis indicates that the structural transition predicted in these hypotheses is developmentally possible (Theißen and Becker, 2004). Among the molecular changes which have been suggested to have caused the modifications in B gene expression are changes in genes that control B gene expression (encoding ‘trans-acting factors’), or changes in the cis-regulatory elements of the B genes themselves (for a more detailed discussion, see Theißen and Becker, 2004).

Under the hypotheses discussed here the perianth originated later than floral bisexuality. A major reason for this assumption is that the identity of all the organs of perianthless flowers could be specified by organ identity genes (class B and class C genes) that were probably established less flowers could be specified by organ identity genes. The assumption is that the identity of all the organs of perianth organisation’ (i.e. bisexuality) in angiosperm flowers might have been established. Mutant analysis indicates that the structural transition predicted in these hypotheses is developmentally possible (Theißen and Becker, 2004). Among the molecular changes which have been suggested to have caused the modifications in B gene expression are changes in genes that control B gene expression (encoding ‘trans-acting factors’), or changes in the cis-regulatory elements of the B genes themselves (for a more detailed discussion, see Theißen and Becker, 2004).

Two alternatives to the out-of-male and out-of-female hypotheses have been inspired by phylogeny reconstructions of the floral meristem identity gene FLORICAULA/LEAFY and its orthologues. The ‘mostly male theory’ also maintains that flower organization derives more from the male structure of ancestral gymnosperms than from the female structure; in contrast to the ‘out-of-male hypothesis’, however, it assumes that flowers originated when female organs (ovules) started to develop ectopically in male cones (Frohlich and Parker, 2000; Frohlich, 2003, 2006). Albert et al. (2002) proposed that ‘sexual condensation’ (i.e. bisexuality) in angiosperm flowers might have originated when previously distinct spatial regulation of copies of LFY-like genes in separate male and female apices was amalgamated into singular LEAF (the proposed orthologue of LEAFY in gymnosperms) control in all reproductive meristems following loss of the NEEDLE gene, a LFY parologue only present in gymnosperms. Albert et al. (2002) hypothesized that positive selection on the partly redundant LEAF parologue trapped reproductively unisexual angiosperm-ancestors into a condensed, bisexual state.

**Why floral quartets?**

All of the hypotheses on flower origin mentioned attribute a major role to class B and class C proteins or to regulators of floral homeotic proteins (LEAFY-like proteins). Even though these transcription factors may well have been of fundamental importance during the origin of the angiosperm flower, we would like to bring to your attention yet another player among the floral homeotic proteins, namely class E proteins. Why? In contrast to the other classes of floral homeotic genes, class E genes are absolutely required for the identity of all floral organs to develop (Ditta et al., 2004). Moreover, in contrast to class B and class C genes (for that matter comprising class D genes), class E genes are unique to flowering plants and appear to be absent in gymnosperms. Attributing class E genes a special role during flower origin thus appears not too far fetched (Zahn et al., 2005).

Hence the question arises as to what makes class E genes so special from a functional, or mechanistic, point of view. A clue may be provided by the floral quartet model; it accurately reflects the developmental importance of the proteins encoded by the class E genes, because class E proteins are constituents of all the predicted multimeric transcription factor complexes (Honma and Goto, 2001; Theißen, 2001; Theißen and Saedler, 2001), and perhaps of others as well (Theißen and Melzer, 2006; Melzer et al., 2006). Why should tetrameric rather than dimeric protein complexes specify organ identity? One could easily imagine a scenario in which all floral organs are determined by only three classes (ABC) of proteins that form specific dimers which then act in a combinatorial way to activate organ-specific target gene expression (see above) (Riechmann et al., 1996).

A ‘null hypothesis’ could postulate that the establishment of tetramer formation represents just a neutral drift in protein–protein interaction patterns without any consequences for gene regulation or the fitness of the affected organisms. We consider this as unlikely, however, and suggest that the transition from a dimeric to a tetrameric mode of protein interaction improved gene regulatory mechanisms by increasing the cooperativity by which MIKC-type MADS-domain proteins bind to DNA.

Cooperative binding describes the phenomenon that binding affinity increases with the amount of molecules already bound. In the case of MADS-domain proteins it means that a dimer bound to a CArG-box increases the likelihood that a nearby CArG-box is also occupied by a MADS-domain protein dimer. Cooperative binding is a common theme among prokaryotic as well as eukaryotic DNA-binding proteins (for examples, see Tsai et al., 1989; Ptashne, 2005; Licht and Brantl, 2006; Svingen and Tonissen, 2006). It offers the (unique) advantage to increase the ratio of specific to non-specific binding while still keeping the system flexible enough for fast and robust regulation (Ptashne, 2004).

In addition, in case of transcription factors binding cooperatively, a ‘genetic switch’ can be created, in that small changes in protein concentration within a certain, relatively small critical range, can cause a dramatic response on target...
genes being regulated. Thus by a small change in the concentrations of transcription factors developmental programmes can be changed dramatically (for discussions on the switch function of cooperative binding proteins, see Ackers et al., 1982; Ptashne, 2004, 2005). Moreover, transcription factors often control their own expression via feedback loops, with this reinforcing a possible switch function.

During developmental transitions as well as organ formation, living cells need regulatory processes of great sensitivity and high specificity, otherwise the transition to reproductive growth may occur at the wrong time or the specification of organ identity may fail. Therefore, signal-to-noise ratios must be high when relevant information is received and transmitted between the cell surface, the cytoplasm and the nucleus (Blundell and Fernández-Recio, 2006). Hence, living cells have complex signalling pathways that are moderated by feedback mechanisms. The transcription factors encoded by floral homeotic genes can be considered as late elements in signal transduction pathways specifying organ identity; there is considerable evidence that they are subject to feedback mechanisms including autoregulatory control (Schwarz-Sommer et al., 1992; Gómez-Mena et al., 2005). Combinatorial tetramer formation by floral quartet formation could well represent an important aspect of the switch and feedback mechanisms that distinguishes the development of one floral organ identity from the others. This may explain the fundamental developmental importance of the function of class E genes. All our – admittedly limited – knowledge about the interaction behaviour of floral homeotic proteins points towards a function of class E (and, to some extent, class A) proteins as mediators of higher order complex formation, i.e. they might constitute the ‘bridge proteins’ that make the floral quartets possible. We assume that without such bridge proteins no higher order complexes will form and, consequently, the system will loose a considerable component of its cooperative behaviour and maybe also (at least partly) its ability to function as a genetic switch. Note, however, that yet another function of SEP-like and AP1-like proteins might be to provide transcriptional activation domains to the corresponding transcription factor complexes (Honma and Goto, 2001).

Baum and Hileman (2006) proposed a model about the role of floral quartets during flower origin that nicely fits to our views. Previously, Theißen and Becker (2004) suggested that during flower origin, changes in cis-regulatory elements, typically located in the promoter region of a respective gene, in enhancers or silencers up- or downstream of the coding region, or even in introns, might have rendered class B gene expression more susceptible to a hypothetical apical–basal gradient already present within the reproductive cone. The authors suggested that such a gradient could be based on either a low molecular weight compound (e.g. a phytohormone) or on a protein (e.g. a transcription factor). Baum and Hileman (2006) suggested that the apical–basal gradient being employed is made up of an orthologue of the floral meristem identity protein LFY, indeed being a transcription factor. In Arabidopsis, LFY is an important signalling integrator for the transition from vegetative to reproductive development. LFY expression appears to respond positively to a developmental clock and to signalling by the phytohormone gibberellin. Baum and Hileman (2006) suggest that reproductive shoots in the most recent common ancestor of flowering plants may have gradually accumulated higher and higher levels of LFY protein during development. Under the hypothesis of Baum and Hileman (2006), both class B and class C genes are activated by LFY-like proteins, but class C genes are more responsive to LFY than class B genes, so that the maximal level of class C gene expression is finally much higher than that of class B gene expression. Therefore, while class C gene expression continues to increase during development due to increased LFY expression, class B gene expression reaches a plateau and then declines because complexes of class C and class E proteins outcompete B/C/E complexes, and hence the autoregulatory circuit of B gene expression is interrupted. Consequently, at the base of the reproductive cone, where both B and C genes are expressed, male organs develop, while at the top, where only C genes are expressed, female organs develop (Baum and Hileman, 2006).

The model of Baum and Hileman (2006) already includes class E proteins and the formation of higher order complexes. To emphasize the evolutionary importance of the class E proteins, we consider the hypothesis of Baum and Hileman (2006) in the light of a hypothetical scenario in which just protein dimers rather than tetramers control floral organ formation (Fig. 2A). We assume that either dimers of class B proteins, together with dimers of class C proteins, and/or B/C heterodimers are responsible for development of male reproductive organs, whereas female organs are specified by class C protein homodimers. Under this scenario the system would lack a considerable component of its cooperative behaviour and might consequently also be a less efficient developmental switch distinguishing male from female development. The critical threshold level of protein concentration needed for organ formation would need to be much higher, because with the loss of some cooperativity DNA-binding affinity would also decrease. Moreover, as class C protein concentration increases along the cone axis, a quite broad zone in which neither unambiguously male nor female organs develop may appear, because the critical threshold concentration for developing female organs is not yet reached, while class C protein level is so high that male organs also do not develop properly (Fig. 2A). According to such a scenario an elongated axis with male reproductive organs on the bottom, followed by a broad zone of ‘unclear organ development’, and finally followed by female organs might develop.

In contrast, if cooperative binding exists (Fig. 2B) DNA-binding affinity would increase. Consequently, protein levels needed for specification of reproductive organs would be much lower (as the DNA-binding affinity is higher, fewer molecules are needed to occupy all the relevant target gene promoters). Also, the concentration of protein complexes bound to target gene promoters along the cone axis would increase with a much steeper
slope and by this constituting the developmental switch. Finally, as a consequence of this cooperative binding, a compressed axis with female organs on the top and male organs on the bottom will develop (Fig. 2B), and the zone of ‘unclear organ formation’ might be considerably smaller. The resulting structure thus resembles an angiosperm flower without a perianth, as predicted in some models of flower origin (see above). We speculate that it is cooperativity, brought about by the floral quartets, that facilitated the development of male and female organs so closely together on a compressed axis. Furthermore, we propose that the many different kinds of organs within a typical angiosperm flower (sepals, petals, stamens, carpels with ovules) can only develop in such close vicinity to each other because each of them is defined by its own genetic switch, i.e. by its own ‘floral quartet’.

One should note that our model considers cooperativity of DNA binding, but not positive autoregulation of class C genes by corresponding proteins, which may further help to define the expression domains of class B and C genes and thus to reduce — or eventually even eliminate — the ‘zone of unclear organ formation’ (Fig. 2).

It is further noteworthy that the increase of cooperative DNA binding might also have increased the spectrum of target genes that can be regulated as the affinity to the respective promoters is increased. This broadening of the target gene spectrum could be another important aspect in the evolution of the angiosperm flower.

Importantly, the scenario presented here is not in contradiction to any of the proposed hypotheses on flower origin outlined above. Rather, it can be understood as an addition to all these theories. The ‘invention’ of floral quartets might have been one of the few crucial molecular steps for the origin of the angiosperm flower. To test this hypothesis, one needs to trace back the origin of the floral quartets, i.e. the origin of the proteins that confer tetramerization. Good candidates would be (phylogenetically defined) SEP- and AP1-like genes, which gave rise to (functionally defined) class E and class A genes, respectively.

How did floral quartets originate?

The quartet (or tetramer) of floral homeotic proteins is actually a dimer of dimers, which provides a first clue as
to how it originated during evolution (Kaufmann et al., 2005). An extreme hypothesis about the origin of floral quartet formation suggests that tetramerization could be an intrinsic capacity of all MIKC-type proteins, based on the unique domain structure of these proteins; if so, quaternary complexes of MIKC-type proteins should exist not only in seed plants, but also in ferns, mosses and even some green algae. Alternatively, quartet formation might be a phenomenon that is restricted to some derived groups of angiosperms, such as core eudicots (and maybe also monocots) (Kaufmann et al., 2005). Pulling together what has been said before, however, we consider another hypothesis more likely – that the capacity of quartet formation was provided by the origin of SEP-like or SE/P-AP1-like genes either before or close to the origin of extant angiosperms. Quartet formation may have improved the developmental switch mechanism underlying the specification of organ identity, as required, e.g. for the sharp transition from female to male (or vice versa) organ identity within one and the same, condensed reproductive structure (‘flower’). So what is known about the origin of the clades of SEP-like and AP1-like genes? (This clade terminology is used for simplicity here, although the correct clade names would be AGL2- and SQUAMOSA-like genes, respectively; see Theißen et al., 1996; Becker and Theißen, 2003.)

As already mentioned above, class B and class C floral homeotic proteins probably originated 300–400 million years ago in the lineage that led to the extant seed plants (Theissen et al., 2000). However, the origin of the gene subfamilies comprising class A and class E genes – AP1-like genes and SEP-like genes, respectively (Theißen et al., 1996; Pelaz et al., 2000) – is unresolved. Despite extensive efforts neither AP1-like nor SEP-like genes have been isolated from conifers, gnetophytes, cycads or Ginkgo, strongly suggesting that these gene types are absent in extant gymnosperms (Becker and Theißen, 2003; Zahn et al., 2005). It is less clear, however, whether AP1- and SEP-like genes originated in the lineage that led to extant angiosperms after the gymnosperm lineage had branched-off, or whether they originated earlier and had been established in the most recent common ancestor of extant seed plants about 300 million years ago already, and then got lost in the lineage that led to extant gymnosperms.

Unfortunately, the phylogenetic relationship among the major clades of MIKC-type genes is poorly resolved, undermining a clear distinction between these two scenarios. AP1- and SEP-like genes consistently form a superclade together with AGL6-like genes (a subfamily of MADS box genes whose function is unknown), but the relationship among these gene clades and to other MIKC-type genes is largely unclear. In quite a number of gene trees AGL6-like genes appear as the sister group to SEP-like genes, to the exclusion of AP1-like genes at the base of the superclade (Fig. 3A; Theissen et al., 2000; Becker and Theißen, 2003; and Nam et al., 2003) with Carlsbecker et al. (2003) and Kim et al. (2005)). The phylogeny shown in (C) is discussed by Becker and Theißen (2003). Dotted lines represent lineages that probably have been lost in the extant gymnosperms. Dimeric and tetrameric complexes bound to DNA are symbolized as in Fig. 2. Asterisks mark the hypothetical origin of higher order complexes. We assume that higher order complex formation was established in the most recent common ancestor of SEP-like and AP1-like proteins.

Fig. 3. Three different scenarios of how SEP-, AP1- and AGL6-like proteins might be phylogenetically related. The phylogenies shown in (A) and (B) are supported by various publications [compare Theissen et al. (2000); Becker and Theißen (2003) and Nam et al. (2003) with Carlsbecker et al. (2003) and Kim et al. (2005)]. The phylogeny shown in (C) is discussed by Becker and Theißen (2003). Dotted lines represent lineages that probably have been lost in the extant gymnosperms. Dimeric and tetrameric complexes bound to DNA are symbolized as in Fig. 2. Asterisks mark the hypothetical origin of higher order complexes. We assume that higher order complex formation was established in the most recent common ancestor of SEP-like and AP1-like proteins.
Kim et al., 2005). If this reflects the true phylogeny, then the gene lineage that later diverged into SEP-like and AP1-like genes in angiosperms must have been lost in the common ancestor of extant gymnosperms (Fig. 3B). A third hypothesis, emphasized by Becker and Theißen (2003), would be that gymnosperm AGL6-like genes are sister to all angiosperm genes, with angiosperm AGL6-like genes being sister to SEP-like and AP1-like genes (Fig. 3C). This would imply that angiosperm AGL6-like genes originated from AGL6-like genes in the common ancestor of extant gymnosperms and angiosperms, and gave rise to AP1-like and SEP-like genes by two consecutive gene duplications followed by sequence divergence. Unfortunately, most phylogeny reconstructions do not corroborate such a scenario so far. However, an even more challenging task is to figure out where and when the developmental functions conferred by these genes have their origin. While the function of class A genes (‘A-function’) is not well conserved even within eudicots – it might just be an offspring of a more widespread function in the specification of meristem identity (Theissen et al., 2000; Litt, 2007) – the opposite is true for the function of class E genes (‘E-function’). Class E gene mutants have been described in eudicots as well as monocots (Pelaz et al., 2000; Ferrario et al., 2003; Agrawal et al. 2005), and the expression patterns of SEP-like genes suggests that there is an E-function already in basally diverging angiosperm lineages (Kim et al., 2005). Due to its fundamental importance in flower development, clarifying the origin of the E-function may well be a major step in understanding the origin of the flower. However, although the A-function, as originally defined – specifying sepals and petals – is poorly conserved, the respective proteins show characteristics that might not only be well conserved, but that are also shared with E-function proteins. Indeed, it is hard to clearly distinguish the function of class A and class E genes, even in eudicot model plants. For example, if AP1 (an A-function gene) is overexpressed together with AP3 and PI (B-function genes), leaves are transformed into petals. However, the same phenotype can be achieved by overexpressing SEP3 (an E-function gene) instead of AP1 (Honma and Goto, 2001). Moreover, recent reports indicate that both, SEP3 and AP1, play a similar role in floral induction (Sridhar et al., 2006). Also, class A and class E proteins appear to be the only floral homeotic proteins that show considerable transcription activation ability (Honma and Goto, 2001). Finally, and maybe most importantly, E- and A-function-related proteins are to date the only ones that have been shown to serve as bridge proteins in higher order complex formation (see Kaufmann et al., 2005, and references cited therein).

The properties shared by A- and E-function proteins (the transcription activation ability, the ability to mediate higher order complex formation) are the ones that are evolutionary most important. One might therefore speak of an E-function that is also shared by A-function-related proteins at the base of extant angiosperms. This is supported by the broad expression pattern of A-genes in basally diverging angiosperm lineages, which more closely resembles the expression of E-genes than that of A-genes in eudicots (Kim et al., 2005).

To trace back the origin of higher order complex formation, both the function and the phylogenetic position of AGL6-like genes, the ‘third player’ in the superclade, needs to be determined. If AGL6-like proteins are nested within SEP- and AP1-like proteins, the most parsimonious assumption would be that higher order complex formation originated once at the base of this superclade and hence also gymnosperm AGL6-like proteins should possess this ability (Fig. 3A). If, on the other hand, AGL6-like genes are sister to both SEP-like and AP1-like genes, one could imagine that the ability to form higher order complexes originated near the base of the SEP/AP1-lineage and is therefore specific for extant angiosperms (Fig. 3B). A similar scenario can be envisaged if the phylogenetic scenario depicted in Fig. 3C is true.

Unfortunately, there are almost no genetic or biochemical data available for AGL6-like proteins that could strongly support either of the views. It therefore remains one of the main goals of future research to examine whether AGL6-like proteins from angiosperms and – maybe more importantly – from gymnosperms have the ability to mediate tetramerization of floral homeotic proteins. This will probably tell us when the floral quartets originated and with this bring us closer to the molecular heart of how the angiosperm flower originated.

FADING BORDERS AND SHIFTING BOUNDARIES: THE ABCS OF FLORAL DIVERSITY

Homeosis and floral diversification

The flowers of the more than 250 000 extant angiosperm species vary in many different ways, including the number, arrangement, identity and shape of floral organs, and the symmetry of the flower as a whole. Arguably the best understood aspect from a developmental genetic point of view is organ identity.

As previously discussed, changes in the expression patterns of floral homeotic genes in transgenic or mutant plants can bring about homeotic transformations of floral organs, so that at some places the wrong organs develop. For example, the expression of class C genes in the whorls of the perianth leads to the development of carpelloid organs in the first floral whorl, where actually sepals should develop; in the second whorl, where petals are usually formed, staminoid organs or even true stamens develop (Bradley et al., 1993). The ectopic expression of class B genes throughout the flower of Arabidopsis results in the development of petaloid organs rather than sepals in the first floral whorl and of staminoid organs rather than carpels in the fourth floral whorl (Krizek and Meyerowitz, 1996). Thus it is tempting to speculate that such changes in organ identity represent suitable models for evolutionary processes that generated some aspects of floral diversity. Homeotic transitions and other very rapid (‘saltational’) morphological changes have long been neglected by mainstream evolutionary biology that considered gradualistic changes the only credible mode of evolution; in the framework of evolutionary developmental biology,
however, homeotic transitions have been reconsidered as a reasonable mode of evolution, and ample evidence has been gathered that they indeed played an important role during the evolution of land plants (Theissen et al., 2000, 2002; Bateman and DiMichele, 2002; Kramer et al., 2003; Ronse De Craene, 2003, 2007; Theissen, 2005b, 2006; Hintz et al., 2006; Nutt et al., 2006).

To determine whether changes in the spatial or temporal expression of floral homeotic genes have contributed to the structural diversification of the flower during angiosperm evolution, comparative expression studies in flowers with different architectures are required. On the following we discuss a few molecular case studies that corroborate the view that homeotic transitions played a role during the evolution of flowers.

Fading borders by fuzzy expression domains

The ‘classical ABC model’ was developed based on studies in core eudicots, so it could well represent a derived situation within angiosperms.; indeed, studies in other groups of angiosperms support that view. The basal lineages of angiosperms represent only a few per cent of the flowering plant species, but it is here where the greatest diversity in floral structure and form is found. The flowers of basal angiosperms vary considerably in size, the number of floral parts and the arrangement of floral organs in spirals or whorls.

The mRNA expression patterns of putative floral homeotic genes (orthologues of the floral organ identity genes from core eudicots) were determined in quite a number of basally diverging angiosperms lineages, including Amborella, Nuphar (Fig. 1), Illicium, Magnolia and Asimia (Kim et al., 2005). Generally, orthologues of class B genes were found to be expressed in stamens and perianth organs, orthologues of class C genes in stamens and carpels, and class E gene orthologues in all floral organs (Kim et al., 2005). For example, water lilies (Nymphaeaceae) such as Nuphar (Fig. 1) have outer floral organs sometimes referred to as sepals because of their greenish color, even though they exhibit expression of class B gene orthologues, as do the petals, staminodes and stamens (Kim et al., 2005). The perianth of Amborella, another basal angiosperm, shows a gradual transition from outer to inner floral organs (lateral receptacular bracts to tepals, from outer tepals to inner tepals, and from tepals to reproductive organs) (Soltis et al., 2006); all these organs express class B gene orthologues (Kim et al., 2005).

Class A and class C gene orthologues have broader expression domains in the floral apices of basal angiosperms than in core eudicots. The expression patterns obtained for ABC genes from basal angiosperms are generally consistent with the morphology of the flowers as predicted by the ABC model. For example, floral organs in Amborella, Nuphar and Illicium expressing class B gene orthologues are always more or less 'petaloid' (or stamens).

These findings gave rise to the 'fading borders model', suggesting that gradual transitions in organ morphology across the floral meristem result from gradients in level of expression of ABC genes across the floral meristem (Fig. 1) (Buzgo et al. 2004; Soltis et al., 2006). Relatively weak expression at the edge of each gene’s expression domain overlaps with expression of another floral homeotic gene, while strong expression of each floral organ identity gene occurs distant from the adjacent gene’s centre of strong expression. According to this hypothesis, in Nuphar, for example, staminodes develop in the area where relatively weak C-gene expression overlaps with strong B-gene expression. As C-gene expression increases, stamens instead of staminodes develop in the succeeding inner whorls (see ABC model for Nuphar in Fig. 1). This fuzzy pattern of ABC gene expression would impose some features of adjacent organs onto each other and thus produce morphologically intergrading rather than distinct floral organs. Such views about ambiguous organ identity may first appear unusual; they have been anticipated, however, by a discussion ongoing for decades about even more radical concepts such as ‘fuzzy morphology’ and ‘continuum morphology’ to explain peculiarities in the body plans of plants; good cases in point are bladderworts (Utricularia) and river-weeds (Podostemaceae) (Rutishauser and Isler, 2001; Rutishauser and Moline, 2005; and references cited therein).

Sharpening the borders by autoregulatory control

If having broad and somewhat fuzzy expression domains is the ancestral condition of floral homeotic genes in angiosperm flowers, the question arises as to how these expression domains became constrained and sharpened during evolution. We suggest that the establishment of positive autoregulatory control contributed significantly to this, because it led to the amplification of small differences in expression levels so that eventually only domains of ‘no expression’ or ‘full expression’ remain, with sharp borders between both expression domains (and hence organs). Autoregulatory control may well have originated when CArG-boxes originated by chance (e.g. point mutations) in the regulatory regions (such as promoters) of floral homeotic genes.

However, also changes in protein–protein interaction patterns and more sophisticated changes in gene regulatory control may have been involved. Class B genes might serve as a well-studied example here. The class B proteins from the eudicot model plants Arabidopsis and Antirrhinum are stable and functional in the cell only as heterodimers of a DEF/AP3-like and a GLO/PI-like protein. These obligate heterodimers regulate their own transcription by binding to CArG-boxes in the promoters of their own genes (or by binding to promoters of positive regulators) and are, therefore, co-expressed in petals and stamens throughout flower development. Heterodimerization is also absolutely required for movement of B proteins into the nucleus and DNA binding in vitro (reviewed by Theißen and Becker, 2004).

In contrast to the obligatory heterodimerization in core eudicots, a diversity of interaction patterns of class B proteins, including obligate and facultative heterodimerization as well as homodimerization, has been found in monocots, such as tulip (Tulipa gesneriana) and lily (Lilium regale) (Kanno et al., 2003; Winter et al., 2002b).
All class B gene orthologues from gymnosperms tested so far revealed the ability to homodimerize (Sundström and Engström, 2002; Winter et al., 2002b).

These data suggest that the interaction of class B proteins evolved from homodimerization in gymnosperms to facultative and obligate heterodimerization in monocots on the one hand, and to exclusively obligate heterodimerization in core eudicots on the other hand (Winter et al., 2002b).

Some types of obligate heterodimerization may have originated in several grasses such as maize (Zea mays) a second time independently from its origin in core eudicots (Whipple and Schmidt, 2006). This ‘trend’ towards obligate heterodimerization raises the question as to whether it is more than a ‘neutral’ drift in protein interaction patterns. One hypothesis maintains that it was exactly the origin of obligate heterodimerization that helped to restrict functional class B gene expression to sharp domains (whorls) within the flower, and thus may have contributed to the evolution of morphologically distinct and unambiguous rather than intergrading floral organs (Winter et al., 2002b).

**Getting diversity by sliding boundaries**

Once sharp borders of ABC gene expression and hence unambiguous organ identities originated, diversification of the flower in terms of floral organ identity was everything but over. One obvious way, already mentioned above, to further diversify floral structures is to change the expression domains of floral homeotic genes and hence to change organ identity in the affected whorls. It seems that nature made rich use of this possibility during evolution. Two different scenarios come to mind: either the expression domain of class B genes changes, or the expression domain of both class A and class C genes changes simultaneously (due to the mutual antagonism between class A and class C genes, at least as revealed in Arabidopsis).

Changes in class B gene expression appear of greater evolutionary importance than changes in class A and class C gene expression. Using the ‘classical ABC model’ as the starting point, two different cases can be distinguished: outward and inward shifts of class B gene expression. Tulip (e.g. *Tulipa gesneriana*), lily (e.g. *Lilium regale*), and most of their relatives in the lily family (Liliaceae), for example, have flowers displaying organ identities quite similar to the ones of higher eudicots. However, first whorl organs are typically petaloid like second whorl organs rather than sepaloid, i.e. the perianth is a perigon composed of two whorls of petaloid tepals (Fig. 1). This suggests a ‘modified ABC model’ that is characterized by an outward shift of class B gene expression, so that it is not only found in the organs of whorls two and three, but also in the organs of the first whorl (Theissen et al., 2000, and literature cited therein). When putative class B genes, i.e. DEF-like and GLO-like genes, were investigated in lily and tulip, they were found to be expressed in the organs of the first three whorls of the flower, as predicted by the modified ABC model (Kanno et al., 2003). A similar situation is realized in orchids (Orchidaceae). Orchid flowers have a unique structure including three types of perianth organs: three outer tepals (T1–T3) in the first floral whorl, and two lateral inner tepals (t1, t2) as well as a median inner tepal (t3) called the lip (or labellum) in the second floral whorl. All these tepals are usually of petaloid appearance and express putative class B floral homeotic genes (DEF-like and GLO-like genes) (Tsai et al., 2004; Xu et al., 2006).

Outside of the monocots, a similar situation is found, for example, in some basal eudicots. Many flowers of the Ranunculaceae have distinctly different petaloid organs in the first two whorls (even though the first whorl organs are usually called ‘sepals’). Kramer et al. (2003) identified many duplication events of putative class B genes at different phylogenetic levels, with *DEF/AP3*-like genes displaying early duplications near the base of eudicots and *GLO/AP3*-like genes more recent duplications. Again, expression studies suggested that petaloidy of organs in the first floral whorl is due to a shift of class B gene expression towards the first floral whorl (as in lily and tulip), but also that differential expression of a particular lineage of *DEF/AP3*-like genes has contributed to the distinction of the petaloid organs in the first and second floral whorl (Kramer et al., 2003).

A somehow complementary case, i.e. an inward shift in class B gene expression, is given in sorrel (*Rumex acetosa*), where only the stamens, but none of the perianth organs express *DEF/AP3*-like or *GLO1/PI*-like (class B) floral homeotic genes; these findings are in line with the sepaloid appearance of first and second whorl organs of the wind-pollinated flowers of sorrel (Fig. 1) (Ainsworth et al., 1995).

These cases strongly suggest that floral diversity can be achieved by outward or inward shifts of class B floral organ identity gene expression (Fig. 1). Respective mechanisms have been suggested by several authors and are now known as ‘sliding boundary’ or ‘shifting boundary’ models (Bowman, 1997; Albert et al., 1998; Theissen et al., 2000; Kramer et al., 2003). Shifts in class A and class C gene expression have been less frequently considered. A plausible possibility would be the outward shift of class C gene expression towards the second floral whorl, thus transforming petals into stamens. Good candidates for flowers specified that way are found in some basal eudicots, such as *Macleaya* (Papaveraceae) (Fig. 1) (Ronse De Craene, 2003) and in some naturally occurring floral homeotic varieties of *Capsella bursa-pastoris* (Brassicaceae), a close relative of Arabidopsis (Hintz et al., 2006; Nutt et al., 2006).

Unfortunately, all the cases outlined above await corroboration by mutant analysis.

**Right on target: different organs by changes in downstream genes**

In some cases very differently appearing organs are specified by very similar ABC systems. An intriguing example is revealed by the comparison between the flowers of grasses (Poaceae, such as *rice, Oryza sativa*), and the flowers of most core eudicots, including Arabidopsis (Fig. 1). While class A genes alone specify sepalas in Arabidopsis, they may specify palea and lemma...
in grasses. Even more remarkably, while the combination of class A and class B genes specifies (usually) showy petals, the same combination of floral homeotic genes specifies lodicules in grasses. Lodicules are grass-specific, small, glandular-like organs that swell at anthesis to spread the lemma and palea apart so that the wind can disperse the pollen produced by the stamens. They are very different from petals indeed. According to their position in the flower, however, lodicules might be petal homologues. Recent comparative expression studies of class B genes in a basal grass and close relatives support that view (Whipple et al., 2007). The development of tremendously different, yet homologous floral organs under the control of orthologous homeotic genes is most plausibly explained by changes in target genes of floral homeotic genes. Future determination and characterization of larger numbers of target genes in both grass and eudicot species than has been possible so far will be required to work out the details of this process.

OUTLOOK

There are quite a number of studies about the importance of gene duplications, changes in protein–protein interactions, or morphological innovations during evolution. Only few investigations, however, combine these fields of inquiry, even though the processes they study often may be intimately linked. For example, according to one of our hypotheses, the origin of the flower may have been facilitated by a transition from dimerization to tetramerization of floral homeotic proteins, which again could have depended on the establishment of SEP-like or API-like genes by gene or genome duplication and sequence divergence. To test such hypotheses, evolutionary biology has to become an even more ‘interdisciplinary’ endeavour, ranging from hard-core physico-chemistry (e.g. of protein–protein and protein–DNA interactions) via molecular evolution, comparative genomics and bioinformatics (e.g. modelling the consequences of changes in protein–DNA and protein–protein interactions), to field ecology (e.g. investigating reproductive fitness in natural environments), all focused on one research topic (such as floral origin or diversification) rather than being scattered. Stay tuned!

ACKNOWLEDGEMENTS

G.T. thanks Mike Jackson and Thomas L. Rost for their invitation to write a review on ‘the molecular origins of the angiosperm flower’, and for their willingness to include it in this ‘Evo–Devo’ Highlight issue of the Annals of Botany. Special thanks to Cristina Ferrandiz and Charlie Scutt for their organizational efforts concerning the Evo-Devo session at the XV FESPB Congress ‘Plants, People, Ecosystems and Applications’, July 2006, Palais des Congrès, Lyon, France. R.M. thanks the Studienstiftung des deutschen Volkes for his fellowship. Both authors thank Hannelore Simon and Guilan Yang for providing pictures of tulip and rice, respectively.

LITERATURE CITED


Krizek BA, Meyerowitz EM. 1996. The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. *Development* 122: 11–22.


