Annual Plant Reviews (2009) **38**, 1–34 www.interscience.wiley.com
doi: 10.1002/9781444314557.ch1

Chapter 1

CARPEL EVOLUTION

Aurélie C.M. Vialette-Guiraud and Charlie P. Scutt

Laboratoire de Reproduction et Développement des Plantes (CNRS UMR 5667-INRA-ENSL-UCBL), Université de Lyon, Lyon Cedex, France

Abstract: The carpel is the progenitor organ to the fruit and a defining feature of the flowering plants, or angiosperms. This organ has evolved in the angiosperms to generate a wide diversity of forms, often related to breeding strategies and seed distribution mechanisms. In this chapter, we focus on a number of key stages in the evolution of the carpel and fruit, about which something can be said of the molecular mechanisms underlying evolutionary change. In particular, we describe hypotheses for the evolutionary origin of the carpel in the first flowering plants and attempt to reconstruct the history of its structural diversification in various major angiosperm groups. In doing so, we concentrate on the genes and mechanisms whose presence can be deduced at key evolutionary stages in the angiosperms, and on molecular-evolutionary processes such as neo- and sub-functionalization, which have moulded these genes and the developmental processes they regulate. We also review the literature on the evolution of syncarpy – a phenomenon of enormous adaptive significance in the angiosperms. Lastly, we describe some examples of convergent evolution that have led to the development of fruit-like structures both within and outside the flowering plants.

Keywords: carpel; fruit; evolution; development; angiosperm; syncarpy

1.1 The importance of having carpels

The carpel is the female reproductive organ that encloses the ovules in the flowering plants, or angiosperms. By contrast, the ovules of the remainder of the seed plants, the gymnosperms, are most frequently naked structures borne in the axils of leaf-like organs. The carpel is thought to confer several major selective advantages on the flowering plants. Firstly, carpels protect, both physically and biochemically, the ovules within them: many classes of carpel-specific genes encode proteins associated with defence against insects or micro-organisms (Scutt *et al*., 2003). Secondly, highly efficient systems have evolved to facilitate pollen capture and pollen tube guidance in angiosperm carpel tissues, which probably represent considerable improvements over

the mechanisms which bring about fertilization in gymnosperms. Thirdly, during the phase of pollen germination and growth, the carpel provides a site for the operation of self- and inter-specific incompatibility mechanisms which may confer important evolutionary advantages. Accordingly, self-incompatibility prevents close inbreeding, while inter-specific incompatibility prevents too wide hybridizations that may lead to infertile offspring. Fourthly, after fertilization, carpel tissues undergo further developmental changes to form fruits, which protect the developing seeds within them and, at maturity, contribute to the dissemination of these by a wide variety of mechanisms.

Carpels and fruits have evolved to generate a tremendous biodiversity of forms in different angiosperm lineages. The novel carpel and fruit structures thus generated are often linked to diversification in factors including pollinators, breeding systems, seed structure and seed dispersal mechanisms. In numerous angiosperm groups, carpels have fused together to form a syncarpic pistil. A syncarpic arrangement provides a single stigmatic surface, giving a common point of access to all the ovules in the flower. Syncarpy may also allow for heavier pollination vectors and larger fruits with more elaborate seed dispersal mechanisms. For all of the above reasons, carpels and fruits were almost certainly of key importance in the evolutionary success of the angiosperms, which arose from an unknown common ancestor living some 160 MYA (million years ago) (Davies *et al.*, 2004) to generate over 300 000 species alive today.

The immense biodiversity of carpel and fruit development in the extant angiosperms means that no thorough or comprehensive treatment of this subject can realistically be undertaken. In this chapter, we will therefore concentrate on a few key stages in the evolution of carpels and fruits, about which something can be said of the molecular mechanisms underlying evolutionary change. Accordingly, we first describe a number of hypotheses for the evolution of the first carpels and fruits in the flowering plant lineage and review the literature on the likely state of the female reproductive structures in the last common ancestor of the extant angiosperms. We then describe the molecular and morphological differences between carpel development which have arisen following speciation events at two key stages in angiosperm evolution: the last common ancestors of the euangiosperms (including monocots and eudicots) and of the core eudicots (including rosids, asterids and Caryophyllales). We particularly concentrate on such molecular-evolutionary processes as sub- and neo-functionalization, as these apply to the genes of carpel and fruit development. As syncarpy represents an evolutionary change within the angiosperms that was clearly of enormous adaptive significance, we also review the literature on this subject, again concentrating on examples on which something can be said of the molecular mechanisms involved. Lastly, we look briefly at fruit-like structures that, to a botanist, are not fruits. Examples of convergent evolution that has generated structures resembling fruits can be found both within and outside the flowering plants.

1.2 Hypotheses of carpel origin

The origin of the carpel is intimately linked to that of the flower – a structure for which several technical definitions have been proposed. For example, Arber and Parkin (1907) suggest 'an ampisporangiate strobilius of determinate growth and with an involucrum of modified bracts'. According to this definition, the flower equates to a single, compacted developmental axis, or strobilus, rather than a multiply branched system in which each floral organ would correspond to a reduced strobilus. The male and female reproductive organs of the flower are separated into two zones: a central gynoecium, containing the carpels, is typically surrounded by an androecium, containing the stamens.

A number of hypotheses have been formulated to account for the unique features of the flower (Fig. 1.1). As our interest here centres on the carpel, we will concentrate on what these hypotheses have to say about that organ type. It should first be mentioned that a hypothesis proposed by Goethe, the German philosopher, poet and dramatist (von Goethe, 1790), which is now well supported by molecular genetic evidence (Honma and Goto, 2001), regards all plant lateral organs, including carpels, as mutually homologous. Such lateral organs, which form on the flanks of the stem apical meristem or floral meristems, can accordingly be regarded as variants of a basic leaf-type developmental ground plan. Though carpels may be homologous to leaves, these reproductive organs are almost certainly more directly related to leaf-like organs in the reproductive structures of gymnosperms. On this subject, the existing hypotheses for flower origin divide conceptually into two types, depending on whether they regard the carpel as derived by the modification of male or female structures in the presumed gymnosperm-like ancestor of the flowering plants.

The Mostly Male Theory (MMT) (Frohlich and Parker, 2000; Frohlich, 2003) postulates the flower to be mainly derived from the male strobili, or male cone-like structures, of a gymnosperm-like ancestor. According to this hypothesis (Fig. 1.1a), the ancestor of the flowering plants would first have generated ectopic ovules on (male) microsporophylls, which would thereby have become bisexual. The MMT postulates that ectopic ovules were concentrated on sporophylls near the apex of the strobilus, and that the sporophylls bearing these ovules subsequently lost their ability to produce microsporangia, thus becoming female. These newly female sporophylls would then have closed around the ovules to form, in effect, the first carpels. In subsequent evolutionary steps, the residual, entirely female strobili of these proto-flowering plants would have been lost, leaving only bisexual reproductive axes containing apical carpels and basal microsporophylls which would later become stamens.

The MMT is based on evidence from a number of sources, including molecular evidence linked to the *LEAFY* (*LFY*) gene, which acts upstream of genes that specify the identities of floral organs in typical angiosperm flowers. In at

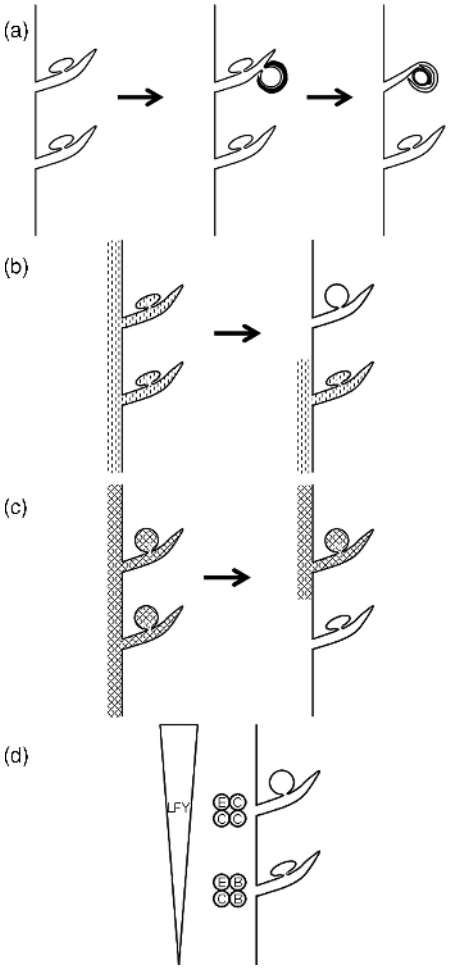

Figure 1.1 Hypotheses for the origin of the flower and its carpel. (a) According to the Mostly Male Theory (Frohlich, 2003), ectopic ovules developed on previously male sporophylls. In a second step, these sporophylls lost their (male) microsporangia and closed around the ovule to form the carpel. The outer integument of the angiosperm ovule (thick line) was formed from a pre-existing female cupule structure. (b) According to the Out-of-Male hypothesis (Theissen *et al.*, 2002), the basipetal movement of male-determining, MADS-box B-function gene expression (shaded area) in a male strobilus left female structures at the apex, which later became carpels. (c) According to the Out-of-Female hypothesis (Theissen *et al.*, 2002), the acropetal movement of MADS-box B-sister gene expression (shaded area) in a female strobilus left male structures at the base, which later became stamens. Female structures at the apex became carpels. (d) According to the Baum and Hileman hypothesis (Baum and Hileman, 2006), a temporal switch in the transcriptional regulation of B- and C-function MADS-box genes by LEAFY (LFY) occurred in an ancestor of the flowering plants. This change generated high concentrations of C-function-rich MADS-box protein complexes at late developmental stages, causing the patterning of the strobilus into apical female and basal male reproductive structures, which later became carpels and stamens, respectively.

least certain gymnosperm taxa, a paralogue of *LFY*, termed *NEEDLY (NLY*), appears to be expressed principally in female cones (Mouradov *et al.*, 1998). The orthologue of *NLY* seems to have been lost from the angiosperm lineage, after its separation from that of the living gymnosperms. The MMT postulates that the loss of *NLY* was accompanied by a more extensive loss of female-specific developmental programmes during the evolution of the flower. Hence, the MMT regards the carpel and the rest of the flower, with the exception of the ovule, as historically male. It should be noted that several studies have brought into question the proposed sex-specific expression of *LFY* and *NLY* in gymnosperms (Carlsbecker *et al.*, 2004; Dornelas and Rodriguez, 2005). Vazquez-Lobo *et al.* (2007) have recently postulated a rather different partitioning of functions between these genes into early and late roles in reproductive development. However, it should also be noted that the hypothesized sex-specific expression of *LFY* and *NLY* in the (unknown) ancestor of the flowering plants is not an absolute requirement for the MMT. Rather, the loss of *NLY* is correlative evidence for the MMT, of which the key postulate is that of ectopic ovule development, which leads to the evolution of carpels from previously male sporophylls.

An attractive feature of the MMT is that it provides an explanation for several unique aspects of reproductive development in angiosperms. For example, this hypothesis is formulated with fossil gymnosperms of the extinct order Corystospermales in mind as potential ancestors of the flowering plants. Corystospermales produced ovules enclosed within cupules, which were borne on unisexual female axes. According to the MMT, such female cupules would have become the outer integument of the ovule – a developmental feature which, like the carpel, is specific to the angiosperms.

Several further hypotheses of flower origin have been proposed, which differ from the MMT in that they postulate the bisexuality of the flower to have arisen by a spatial or temporal change in factors governing the sex of reproductive organs. Hence, these hypotheses do not, in contrast to the MMT, postulate the extensive loss of female developmental programmes during angiosperm evolution. It follows that these various hypotheses would regard the carpel as homologous to the female reproductive structures in gymnosperms. The Out-of-Male (OOM) hypothesis (Theissen *et al.*, 2002) proposes the bisexual flower to have evolved by the basipetal movement of male-promoting, B-class MADS-box gene expression in a previously male strobilus, leaving female structures at the apex (Fig. 1.1b). A sister hypothesis to the OOM hypothesis, the Out-of-Female (OOF) hypothesis (Theissen *et al.*, 2002), postulates a sex-determining role for B-sister MADS-box genes, whose expression is proposed to have moved acropetally in a female strobilus to leave male structures in basal positions (Fig. 1.1c). The identification of the function of a B-sister gene in *Arabidopsis thaliana*, which proves to determine coloration in the outer integument (Nesi *et al.*, 2002), has hardly provided support for the OOF variant of the above two hypotheses. However, it is certainly possible that B-sister genes played a much more central role in

female organ identity in early stages in the evolution of the flowering plants, and may continue to do so in basal flowering plant lineages. In general, therefore, the OOM and OOF hypotheses look below the level of *LFY* and *NLY* in the hierarchical control of gene expression, and postulate a spatial change in MADS-box gene expression, forming a boundary of B- or B-sister class gene expression in a previously unisexual strobilus, which would thereby have become bisexual.

Baum and Hileman (2006) have formulated a further, nameless, hypothesis (we will call it the B&H hypothesis), to account for the evolution of the first flowers (Fig. 1.1d). Like the MMT, the B&H hypothesis proposes a central role for *LFY* in the origin of the flower, but postulates that the origin of floral bisexuality was caused, not by the loss of female-specific developmental programmes, but by a temporally generated switch in the response to *LFY*. According to this hypothesis, LFY protein builds up over time in the meristems of developing reproductive axes and, at a certain threshold of LFY concentration, these meristems switch from the production of (male) microsporoplylls to (female) megasporophylls. This hypothesized switch may involve the action of *LFY* cofactors, such as the ancestors of the *Arabidopsis* F-box protein UNUSUAL FLORAL ORGANS (UFO), and transcription factor WUSCHEL (WUS). Whatever the precise mechanism, the B&H hypothesis proposes that a difference occurred in the relative response to *LFY* of B- and C-class MADS genes during early flower evolution. Accordingly, C-class proteins are proposed to have predominated at high *LFY* concentrations, encountered at the apex of the strobilus at late developmental stages, resulting in MADS-box complexes that were rich in C-class proteins. These proteins would have formed C-rich complexes which would then have specified the development of megasporophylls at the apex of the strobilus.

The above hypotheses are, to some extent, testable. Baum and Hileman (2006), for example, propose a list of predictions that could be tested in basal angiosperm and gymnosperm lineages to support or refute their hypothesis. The MMT stands out from the other hypotheses in proposing the extensive loss of female developmental programmes during early flower evolution. This prediction might provide a means to eliminate either the MMT, or the other contending hypotheses, from consideration. Essentially, if the MMT were correct, we might expect to find numerous classes of genes with female-specific expression patterns in gymnosperms, whose orthologues have apparently been lost from the angiosperm lineage. Gymnosperm genes with male-specific expression patterns should not be affected in this way. The full-scale testing of the MMT by this method has yet to be performed. However, one question mark concerning such a test relates to the degree to which male and female developmental programmes in gymnosperms might be based on different sets of genes, rather than on subtle differences in the expression patterns of a common set of genes. If the latter is predominately the case, this relatively simple method of hypothesis testing may be unavailable.

All of the hypotheses discussed above for the evolution of the flower concentrate to a large extent on the origin of bisexuality, with only the MMT

explicitly accounting for the origin of the carpel. Additionally, it must be admitted that even the MMT does not go into detail on the molecular mechanism of carpel closure. Thus, even if substantial evidence were to favour one of the above hypotheses, elevating it to the status of a theory, many of the mechanistic gaps would still have to be filled in concerning carpel evolution.

1.3 A phylogenetic framework for studies of carpel evolution

Evolutionary studies of the carpel and other unique features of the flowering plants are possible in part due to the contribution of molecular data to phylogenetic analyses of the flowering plants and their relatives. Studies in the 1990s (Goremykin *et al.*, 1996; Winter *et al.*, 1999) indicated the extant seed plants to divide into the two monophyletic groups of the extant angiosperms and gymnosperms, respectively (Fig. 1.2). These findings replaced an earlier hypothesis, based on morphological data, that the small gymnosperm order Gnetales might form a sister group to the flowering plants, as discussed by Donoghue and Doyle (2000). Rather, it would seem that the living gymnosperms, including Gnetales, are monophyletic and share a more distant common ancestor with the flowering plant lineage, which would have lived some 300 MYA (Savard *et al.*, 1994; Goremykin *et al.*, 1997). According to the current consensus view of seed plant phylogeny, the living gymnosperms can be divided into five groups, with Cycadales in the most basal position, followed by the monotypic Ginkgoales, represented only by *Ginkgo bioloba.* In the crown group of living gymnosperms, most molecular phylogenetic studies split the conifers into two, placing Pinaceae in a sister position to Gnetales, as discussed by Kuzoff and Gasser (2000). This split leaves a clade of remaining conifers (Fig. 1.2) that, following the classification of Page (1990), would be composed of Araucariaceae, Cephalotaxaceae, Cupressaceae, Phyllocladaceae, Podocarpaceae, Sciadopityaceae, Taxaceae and Taxodiaceae.

Molecular phylogenetic analyses have also clearly identified the first-diverging lineages within the angiosperm clade, as reviewed by Kuzoff and Gasser (2000). According to these studies, three extant lineages, Amborellales, Nymphaeales and Austrobaileyales, collectively known as the ANA grade, would have diverged from a remaining common lineage at an early stage in the evolution of the flowering plants (Fig. 1.2). Amborellales contains the single species *Amborella trichopoda,* a small tree which is endemic to the tropical island of New Caledonia in the Southern Pacific. Nymphaeales contains three families of herbaceous aquatic plants: Nymphaeaceae, Cabombaceae, and the recently added Hydatellaceae (Saarela *et al.*, 2007). Austrobaileyales contains the four families of Austrobaileyaceae, Illiciaceae, Schissandraceae and Trimeniaceae, which are mostly shrubs, climbers or small trees. There is good evidence, both from phylogenetic analyses and from INDEL (insertion/deletion) mutations (Aoki *et al.*, 2004; Stellari *et al.*, 2004), that Amborellales

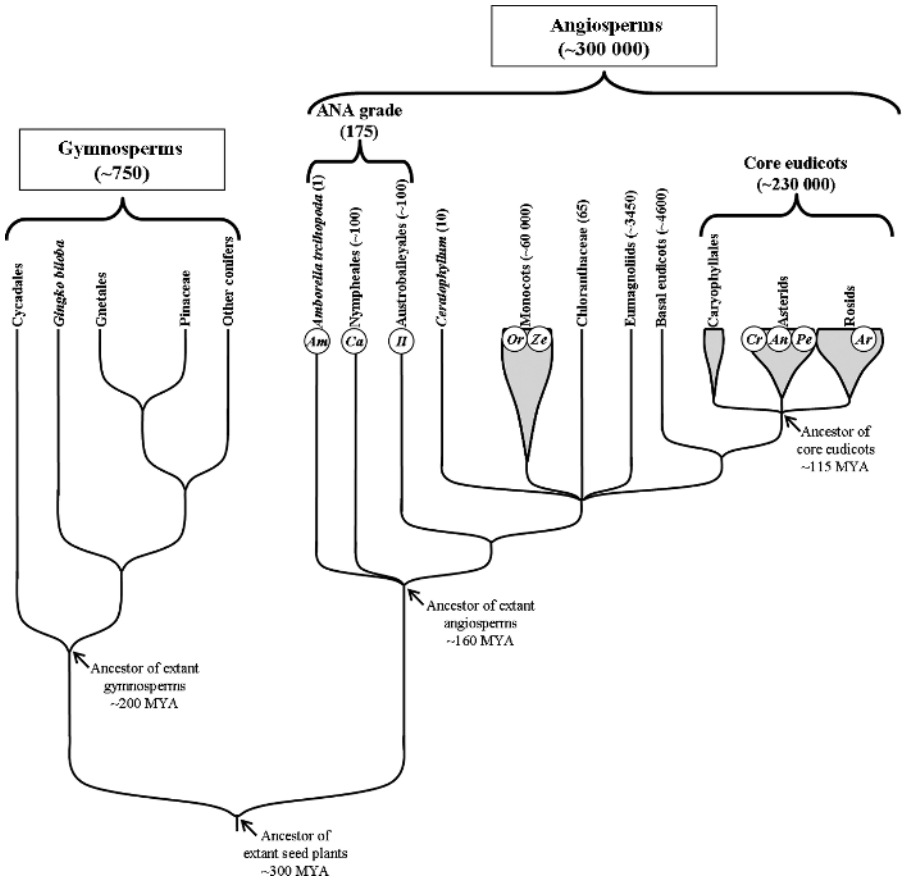

Figure 1.2 The phylogeny of the seed plants, based on a consensus of molecular phylogenetic studies. The numbers of species in major clades are given in parentheses, while approximate dates of divergence are taken from Davies *et al.* (2004), based on a calibration of the molecular clock using fossil data. The positions of some of the taxa referred to in the text are indicated as follows: *Am*, *Amborella trichopoda*; *An*, *Antirrhinum majus*; *Ar*, *Arabidopsis thaliana*; *Ca*, *Cabomba*; *Cr*, *Catharanthus roseus*; *Il*, *Illicium*; *Ze*, *Zea mays* (maize); *Pe*, *Petunia hybrida*; *Or*, *Oryza sativa* (rice).

and Nymphaeales diverged from the remaining angiosperm lineage before Austrobaileyales. However, the relative order of divergence of Amborellales and Nymphaeales remains unclear. Most molecular phylogenies have placed Amborellales alone in the most basal position (e.g. Zanis *et al.*, 2002), while others have grouped it together with Nymphaeales in a first-diverging clade (Qiu *et al.*, 2001).

The remaining angiosperm lineage, after the divergence of the ANA grade, diversified to give five groups with living representatives: *Cerratophyllum*, Chloranthanceae, eumagnoliids, monocots and eudicots (Fig. 1.2). Of these,

the eudicots and monocots together account for over 95% of the estimated 300 000 extant angiosperm species, while some 6500 species of eumagnolids are known (Davies *et al.*, 2004). The short internal branches connecting the five groups of euangiosperms, compared to their long terminal branches, mean that even very large DNA data sets have failed to convincingly resolve their relative points of divergence within the flowering plants (Moore *et al.*, 2007).

Within the eudicots, molecular phylogenetic analyses have clearly identified a number of early-diverging lineages, of which Ranunculales occupies the most basal position. The divergence of these basal lineages leaves a crown group eudicots that includes the major clades of the rosids, asterids and Caryophyllales (Fig. 1.2). The last common ancestor of these 'core eudicots' seems to have been a critical stage in angiosperm evolution, perhaps of equal evolutionary significance as the origin of the flowering plants itself. Indeed, many of the genes controlling flower development in model eudicots seem to be derived from a large-scale or whole-genome duplication that occurred in an ancestor of the core eudicots (Litt and Irish, 2003; Vandenbussche *et al.*, 2003; Kramer and Hall, 2005).

Phylogenetic studies do not of themselves explain the origin of the flower and its carpel. However, these studies provide an essential framework in which we may attempt to answer these questions, which Charles Darwin famously described as an 'Abominable Mystery'. Part of the mystery surrounding the origin of the flower is the evolutionary distance of the flowering plants from their nearest living relatives, the extant gymnosperms, with no continuum of intermediate forms known from the fossil record. Molecular clock estimates, calibrated from well-documented fossil divergences within the angiosperms, suggest the last common ancestor of this group to have lived some 160 MYA (Davies *et al.*, 2004). This date is in reasonable agreement with the first appearance of angiosperm fossils, corresponding to the Lower Cretaceous period, some 130 MYA (Friis *et al.*, 2005). We must therefore conclude that the origin of the carpel, among other unique angiosperm features, occurred some 140 MY after the last common ancestor of the seed plants.

It seems likely that the flower resulted from a combination of molecular and ecological or environmental factors. On the molecular level, a large-scale gene or whole-genome duplication seems to have preceded the radiation of the angiosperms (De Bodt *et al.*, 2005), from which many pairs of paralogous genes that function in reproductive development have been retained. This hypothesized duplication may therefore have provided the raw material for neo-functionalization events on a large scale, which may have been necessary to generate such a novel structure as the flower. However, the generation of polyploids, corresponding to whole-genome duplications, is relatively common in plants (much more so than in animals). It therefore seems probable that the appearance of flowering plants in the Late Jurassic or Early Cretaceous, rather than at any earlier time since the last common ancestor of the seed plants, occurred not only because of a whole-genome duplication, but also in response to specific ecological or environmental conditions. Among such

ecological factors, co-evolution with novel groups of insects that acquired roles as pollinators seems likely to be of significance (Barrett and Willis, 2001).

Certain novel features that contributed to the success of the early flowering plants may have been external to the flower itself. It has been proposed, for example, that the first flowering plants were shrubs or small trees, with a short lifespan compared to their gymnosperm-like ancestors (Feild *et al.*, 2003, 2004). These somewhat ephemeral species are proposed to have found a successful ecological niche in disturbed soils along shaded riverbanks. Other suggestions for ecological factors that may have helped select for the flowering plants include a change in the feeding behaviour of herbivorous dinosaurs, related to a change in their jaw anatomy, and a change in atmospheric carbon dioxide concentration, as discussed by Barrett and Willis (2001). These conjectures serve to illustrate that the fullest possible explanation for the origin of the angiosperms will be achieved only by combining molecular data with the results of morphological and ecological studies and, if possible, with key evidence from fossil groups. In the absence of a living, close relative to the flowering plants, at least some novel fossil evidence would seem to be essential if we are to fill in the exasperating gap of 140 MY of evolution which preceded the origin of the flower.

1.4 A morphological portrait of the ancestral carpel

Morphological comparisons of ANA grade angiosperms (Fig. 1.3a–f) have enabled a number of conclusions to be made on the likely state of the flower, and of the carpel, in the last common ancestor of the living flowering plants (Endress and Igersheim, 2000; Endress, 2001). Accordingly, the flowers of this ancestral species were probably small, bisexual and protogynous. Its carpels were likely to have been simple (apocarpic), rather than fused together into a syncarpic pistil. The stigmatic tissues that permitted the capture of pollen grains in the angiosperms' ancestor were likely to have been covered in multicellular protrusions and would probably have secreted a sticky liquid to hold and supply water to pollen grains during germination. Pollen tubes would then have grown towards the ovary through a canal or aperture in the carpel, which would have contained substances secreted from the carpel margins. Self-incompatibility (SI) systems operating between female tissues and pollen grains are present in some ANA angiosperms, including *Austrobaileya scandens* (Prakash and Alexander, 1984) and *Trimenia moorei* (Bernhardt *et al.*, 2003). However, it is not yet clear whether these SI systems in distantly related Austrobaileyales are homologous, and still less certain that such a system would be ancestral in the entire angiosperm clade. Interestingly, *Amborella*, the only representative of the likely most basally diverging angiosperm lineage, Amborellales, avoids inbreeding by dieocy, rather than through an SI mechanism. However, female *Amborella* flowers contain a nonfunctional stamen, or staminode, which would seem to indicate *Amborella* to be descended from a bisexual ancestor.

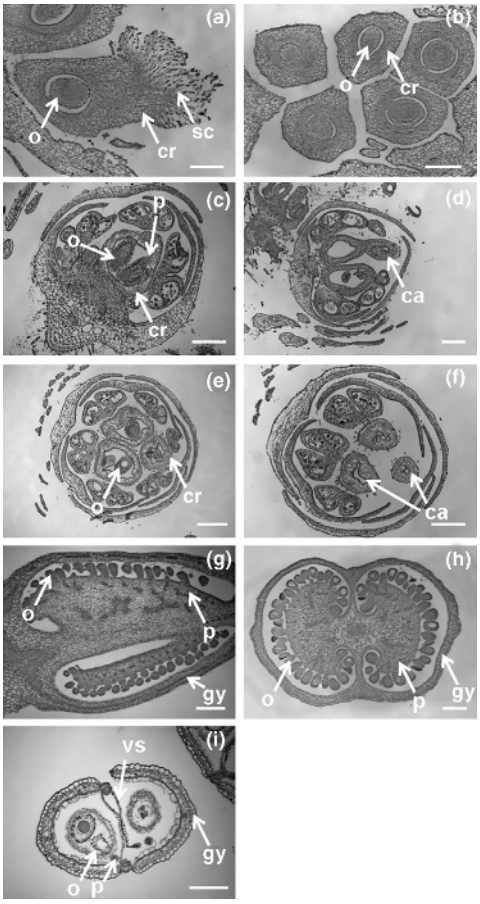

Figure 1.3 A comparison of carpel structures in ANA grade and core eudicot taxa. (a) Longitudinal section of an *Amborella trichopoda* (ANA grade, Amborellales, Amborellaceae) carpel (cr) showing its single anatropous ovule (o) and stigmatic crest (sc) of ridged tissue, which harbours an aperture through which pollen tubes may grow. (b) Transverse section of an *A. trichopoda* female flower bud showing the five separate carpels (cr) of the apocarpic gynoecium. (c) Longitudinal section of a *Cabomba aquatica* (ANA grade, Nymphaeales, Cabombaceae) flower bud showing two anatropous ovules (o) attached to the placenta (p) of one of the three carpels (cr) present. (d) Longitudinal (slightly oblique) section of a *C. aquatica* flower bud showing the secretion-filled canal (ca) in the style, through which pollen tubes may grow. (e) Transverse section of a *C. aquatica* flower bud showing the ovary tissues of the three separate carpels (cr). One or both ovules (o) are visible in each carpel. (f) Transverse section of a *C. aquatica* flower bud showing the secretion-filled canal (ca) in the style of each carpel. (g) Longitudinal section of the syncarpic gynoecium (gy) of *Petunia hybrida* (core eudicots, asterids, Solanaceae) showing many ovules (o) attached to an axile placenta (p). (h) Transverse section of the *P. hybrida* gynoecium showing its two fused carpels, axile placentation (p) and many ovules (o). (i) Transverse section of the syncarpic gynoecium (gy) of *Arabidopsis thaliana* (core eudicots, rosids, Brassicaceae), which is divided into two locules by the post-genital development of a vertical septum (vs). Placentation (p) is parietal. All scale bar represent 250 μm.

The carpel of the ancestral angiosperm probably contained a single ovule of anatropous placentation. This ovule would probably have been covered by two integuments that enclosed an embryo sac and a large nucellus (for more on ovule development, see Chapter 3 of this issue). Double fertilization, leading to the production of an embryo and a bi-parental endosperm, also appears to be a pleisiomorphic feature of the angiosperms. However, the likely cellular arrangement of the embryo sac in the last common ancestor of the angiosperms, and the ploidy of its endosperm tissue, remain open to question. In the majority of flowering plants, the embryo sac arrangement is of the *Polygonum* type, which contains seven cells, one of which, the central cell, is binucleate (Fahn, 1982). The two nuclei of the central cell combine with one sperm nucleus on fertilization of the *Polygonum*-type embryo sac to generate a triploid endosperm. In *Nuphar* (Williams and Friedman, 2002) and *Hydatella* (Friedman, 2008) of Nymphaeales and in *Illicium* (Williams and Friedman, 2004) of Austrobaileyales, the embryo sac contains only four cells, including a uninucleate central cell. Double fertilization in these ANA grade species generates a diploid embryo and a diploid, rather than a triploid, endosperm. However, studies of *Amborella*, which represents the likely most basally diverging ANA grade lineage, indicate a different embryo sac arrangement. The *Amborella* embryo sac contains eight cells, including a binucleate central cell that produces a triploid endosperm after fertilization (Friedman, 2006). The extra cell in the *Amborella* embryo sac, by comparison to the *Polygonum* type, is in the egg apparatus, which thus contains four cells in *Amborella.*

Interestingly, *Hydatella* produces a perisperm, or embryo-nourishing tissue derived from maternal cells, which develops significantly prior to fertilization (Friedman, 2008). This feature resembles the arrangement in gymnosperms, in which entirely maternal tissues fulfil the role of nourishing the embryo, and in which substantial reserves are laid down before fertilization is effected. In the absence of a clear conclusion on the embryo sac arrangement in the last common ancestor of the extant angiosperms, the multiplicity of arrangements in ANA grade angiosperms has been interpreted as a sign of early diversification, prior to the selection of features which became standard in the majority of flowering plant groups (Friedman, 2006).

1.5 The genetic control of carpel development in the first flowering plants

Using molecular techniques to compare ANA grade angiosperms with model plants, we can now begin to describe the mechanisms likely to have controlled carpel development in the ancestor of the living flowering plants. Carpel development in the model taxa *Arabidopsis* and *Antirrhinum* is specified, according to the ABC model of flower development, through the expression of C-function MADS-box transcription factors, in the absence of B-function factors (Coen and Meyerowitz, 1991). The molecular genetic ABC model has more

recently been extended to an ABCE model, in which E-function MADS-box factors form various higher order complexes with those of the A-, B- and C-functions (Pelaz *et al.*, 2000; Honma and Goto, 2001; Theissen and Saedler, 2001). According to this model, carpel development is specified by tetramers containing two molecules of E-function, and two molecules of C-function, MADS-box transcription factors (Fig. 1.4a). The unisexual reproductive structures of gymnosperms appear to be specified by the expression of orthologues of B- and C-clade MADS-box genes (Tandre *et al.*, 1995; Becker *et al.*, 2000; Jager *et al.*, 2003; Zhang *et al.*, 2004), certain of which show similar activities to their *Arabidopsis* orthologues in transgenic *Arabidopsis* plants (Tandre *et al.*, 1998; Winter *et al.*, 2002; Zhang *et al.*, 2004) (Fig. 1.4b). These data strongly suggest that some basic elements of the regulation of flower development have been conserved since the last common ancestor of the living seed plants, well before the origin of the flower.

Phylogenetic analyses of the MADS-box family in ANA grade angiosperms and in gymnosperms clearly indicate that a duplication event took place in the C-function MADS-box lineage prior to the last common ancestor of the living flowering plants (Kim *et al.*, 2005). As a result of this duplication, the ancestors of the clade-defining genes *AGAMOUS (AG) from Arabidopsis thaliana*, and *FLORAL BINDING PROTEIN7 (FBP7)* from *Petunia hybrida* (reviewed by Kramer *et al.*, 2004) were generated. The *AG* clade contains angiosperm C-function genes, whereas the *FBP7* clade contains genes involved in ovule development in diverse taxa, including *Petunia, Arabidopsis* and rice. The role of *FBP7*-like genes in ovule development has been defined as the D-function (Angenent *et al.*, 1995; Colombo *et al.*, 1995). This function is postulated to be necessary for ovule development, and its inactivation leads to supernumerary carpels that develop ectopically in the place of ovules. Interestingly, the *FBP7* (D-function) clade appears to have been lost from Ranunculales (Kramer *et al.*, 2004) and, as will be discussed later, has evolved to share its ovule development function with genes of the *AG* clade in some eudicots, including *Arabidopsis.* Both of these observations suggest a degree of functional fluidity between MADS-box genes of the related C- and D-clades. In addition to the C- and D-functions, two clades of *SEPALLATA (SEP)* genes, which encode E-function MADS-box proteins, have been found in basal angiosperms. The genes *SEP1, SEP2* and *SEP4* from *Arabidopsis* appear to be orthologous to one of these ANA grade *SEP*-clades, while *SEP3* appears to be orthologous to the other (Zahn *et al.*, 2005).

The expression of C-function genes in ANA grade angiosperms is mostly limited to the third and fourth floral whorls, while the E-function genes of these species are expressed in all floral organs (Kim *et al.*, 2005) (Fig. 1.4c). These expression patterns closely resemble those of the corresponding genes in *Arabidopsis*, suggesting important elements of the control of carpel identity to have been conserved in distinct lineages throughout angiosperm evolution. Kim *et al.* (2005) did however note some expression of C-function genes in the perianth organs of the ANA grade angiosperms *Amborella* (Amborellales)

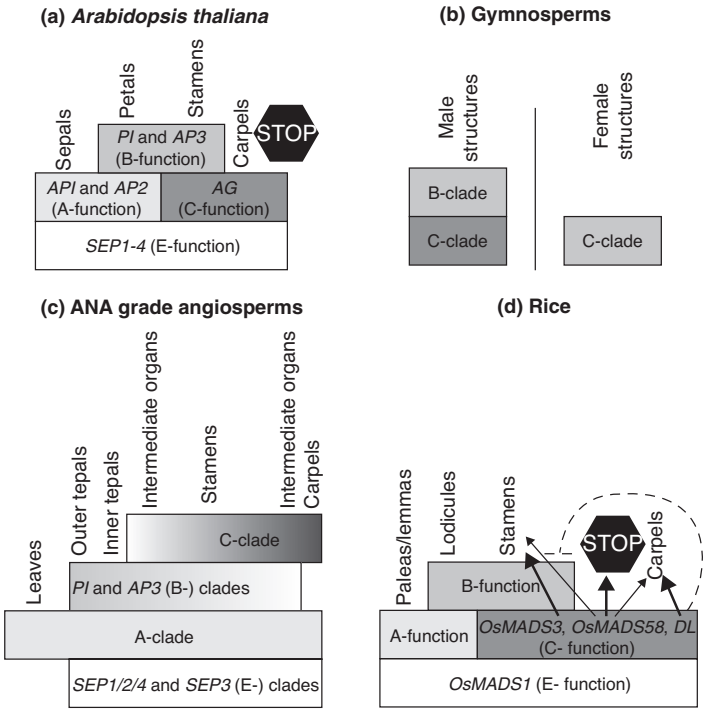

Figure 1.4 The ABCE model of flower development in *Arabidopsis*, and its derivatives in other taxa. (a) In *Arabidopsis*, A-, B-, C- and E-function floral homeotic genes, expressed in overlapping domains (horizontal bars) of the floral meristem, control the identities of floral organs in a combinatorial manner: A + E specifies sepal development in the first whorl, A + B + E specifies petal development in the second whorl, B + C + E specifies stamen development in the third whorl, and C + E specifies carpel development in the fourth whorl. In addition, the C-function causes an arrest of organ proliferation (the 'STOP' function) in the fourth whorl. *AG*, *AGAMOUS*; *AP1*, *APETALLA1*; *AP2*, *APETALA2*; *AP3*, *APETALA3*; *PI*, *PISTILLATA*; *SEP1–4*, *SEPALLATA1–4*. (b) In gymnosperms, B- and C-clade MADS-box genes are expressed in a combinatorial manner in male (B + C) and female (C alone) reproductive structures, resembling the expression of their *Arabidopsus* orthologues in male and female floral organs. (c) In ANA grade angiosperms, B- and C-clade MADS-box gene expression resembles that of the respective *Arabidopsis* orthologues, though with less well-defined boundaries. Strong B-clade gene expression is generally detected in the outer floral whorl of ANA grade angiosperms, possibly reflecting an absence of developmental differentiation between whorls 1 and 2. A-clade MADS-box gene expression differs radically between ANA grade angiosperms and *Arabidopsis*, extending throughout the flower and into leaves. (d) In rice flowers, A-, B- and E-function genes are expressed in similar patterns to those of their *Arabidopsis* orthologues to specify specialized perianth organs (paleas, lemmas and lodicules) and stamens. Two paralogous C-clade MADS-box genes show a partial sub-functionalization between the third and fourth whorls, with one paralogue playing a major role in stamen development in the third whorl, while the other plays a major role in the 'stop' function in the fourth whorl (thick arrows, major roles; thin arrows, minor roles). The YABBY gene *DROOPING LEAF* (*DL)* plays a major role in carpel development that is independent of C-clade MADS-box gene expression. *DL* may act directly on carpel development (solid arrow), or indirectly by limiting the inner boundary of B-function gene expression (dashed arrow), or both of these.

and *Illicium* (Austrobaileyales), in contrast to the more strictly delimited expression patterns of C-function genes in model eudicots. As pointed out by these authors, this observation may reflect the rather gradual transition of floral organ types that is apparent in ANA grade angiosperms, rather than a more fundamental difference in the regulation of carpel development between early and later diverging plant lineages.

In addition to MADS-box genes, the expression patterns of several YABBY transcription factors have been analyzed in ANA grade angiosperms. In *Arabidopsis,* these factors participate in the specification of abaxial cellular identity in lateral organs by defining the side of these organs that faces away from the developmental axis (Bowman, 2000). *CRABS CLAW (CRC)* is a YABBY gene that is expressed only in the abaxial tissues of the gynoecium and in the nectaries of *Arabidopsis* flowers (Bowman and Smyth, 1999). *AmbCRC,* a putative *CRC* orthologue from *Amborella,* shows a similar pattern of expression in carpels to that of *CRC* in *Arabidopsis* (Fourquin *et al*., 2005), suggesting these genes to have conserved a common developmental role since the speciation event that separated their lineages at the base of the flowering plants. *CRC* is a direct target of *AG* in *Arabidopsis* (Gomez-Mena *et al*., 2005), though it is not yet known whether such a direct control relationship exists between the *Amborella* orthologues of these two genes.

INNER NO OUTER (INO) represents a further YABBY gene with a very specific role in female reproductive development in *Arabidopsis*. *INO* is expressed in the outer ovule integument, and its inactivation causes the loss of this angiosperm-specific tissue (Villanneva *et al*., 1999). A putative *INO* orthologue from the ANA grade angiosperm *Nymphaea alba* is expressed in both ovule integuments and in the suspensor (Yamada *et al*., 2003). The broadly similar expression patterns of *INO* orthologues between the ANA grade angiosperm *Nymphaea* and the eudicot *Arabidopsis* suggest the conservation of a role in integument development since the last common ancestor of the flowering plants.

CRC and *INO* are unusual in the YABBY family in showing very specific expression profiles: the other members of this family in *Arabidopsis* are more generally expressed in the abaxial zone of both vegetative and reproductive plant lateral organs (Bowman, 2000). The carpel and outer integument, in which *CRC* and *INO* are respectively expressed, represent pleisiomorphic features of the angiosperms. The relationships of these structures to reproductive organs in gymnosperms could therefore be highly informative of the mechanism by which the flower evolved. Hence, it would be extremely interesting to know whether *CRC* and *INO* orthologues exist in gymnosperms, and if so, to determine their exact expression patterns.

In general, the search for carpel development genes in ANA grade angiosperms has highlighted several instances of the broad conservation of gene functions since the common ancestor of the last flowering plants. In particular, mechanisms involving C- and E-function genes, that specify carpel development at a high level in gene hierarchies, seem to be conserved. However,

much work remains to be done in this field. Many families of transcription factors of known importance to carpel development in *Arabidopsis,* including the auxin response factor (ARF), basic Helix–Loop–Helix (bHLH), MYB and NAC families, have yet to be examined in any detail in ANA grade angiosperms. As major elements of the ABCE model appear to be conserved between angiosperms and gymnosperms, it is at a lower level in regulatory hierarchies, perhaps involving the families listed above, that we might discover the molecular changes that were responsible for the evolution of the carpel in early flowering plants.

EST resources are proving very useful for the analysis of flower development orthologues in numerous seed plant groups (Brenner *et al.*, 2003; Albert *et al.*, 2005; Brenner *et al.*, 2005; Pavy *et al.*, 2005, 2007). However, in any given investigation, the use of such database resources must frequently be complemented by onerous approaches such as cDNA library screening and reverse-transcriptase PCR. With the advent of novel, high-throughput sequencing technologies, the sequencing of the *Amborella* genome from the ANA grade, which has a c-value of 0.89 pg (Leitch and Hanson, 2002), has recently been proposed (Soltis *et al.*, 2008) (The c-value indicates the mass of DNA in one haploid chromosome complement and is thus a convenient measure of genome size.) The sequencing of the *Amborella* genome and other genomes from the ANA grade would certainly provide an invaluable resource for studies of the origin of the flowering plants. The complete sequences of one or more gymnosperm genomes would also be extremely useful in this regard, though this seems still some way off due to the large genome sizes of this group. Gymnosperm genomes range in size from the relatively modest (for a gymnosperm) c-value of 2.25 pg for *Gnetum ula* (Ohri and Khoshoo, 1986) to the very large value of 36.00 pg for *Pinus ayacahuite* (Grotkopp *et al.*, 2004). In particular, the cycads, which are of special interest for their basal position in the gymnosperms, all have large genomes. For example, *Cycas revoluta* has a c-value of 12.75 pg (Ohri and Khoshoo, 1986), some 80-fold larger than that of *Arabidipsis thaliana* (Bennett *et al.*, 2003).

1.6 A major role for the E-function in the origin of the carpel?

E-function MADS-box genes play fundamentally important roles in flower development in *Arabidopsis.* The transcription factors encoded by these genes are hypothesized to act together with combinations of A-, B- and C-function proteins in quaternary complexes that specify the type of floral organ that develops in each whorl of the flower (Pelaz *et al.*, 2000; Honma and Goto, 2001; Theissen and Saedler, 2001). As discussed above, two E-function clades seem to have been present in the last common ancestor of the extant flowering plants (Zahn *et al.*, 2005), perhaps not very long after the origin of the flower. Interestingly however, no *SEP*-like genes have been found in gymnosperms

(Becker and Theissen, 2003). It might therefore seem reasonable to postulate that the origin of the E-function, leading to the generation of quaternary MADS-box complexes, was of fundamental importance for the origin of the flower. Theissen and Melzer (2007) have discussed the possibility that, before the flower, dimers of C-function genes may have specified the development of female reproductive organs, and that the evolution of quaternary MADS-box protein complexes, incorporating both C- and E-function proteins, may thus have represented a key step in the evolution of the carpel. More precisely, the evolution of such quaternary complexes is hypothesized to have caused transcription factor binding to two distinct MADS-box binding motifs, which are termed CArG boxes, in the cis-acting control regions of their downstream target genes. According to this hypothesis, the newly evolved binding behaviour of quaternary MADS-box complexes, also involving A- and B-function proteins in different combinations, would have generated the necessary multiplicity of interactions to specify at least three organ types in early flowers, carpels, stamens and tepals, with the possible later division of tepals into distinct whorls of petals and sepals. Quaternary complexes might also have led to a positive cooperativity of binding to multiple sites in target gene promoters, which might in turn have generated a steeper gradient of transcriptional response between MADS-box genes and their direct targets. Such a steep gradient of response could have generated sharper transitions between the reproductive organs and produced the concomitant compaction of the axis that is apparent in the angiosperm flower.

In discussing the evolutionary role of the E-function, it should be noted, however, that the origin of this function is, like that of the angiosperms, rather mysterious. The four *SEP* genes of *Arabidopsis* form an E-function clade that is closely related to the *Arabidopsis* gene *AGAMOUS-LIKE6 (AGL6)*. Orthologues of *AGL6*, rather than those of the combined *AGL6* + *SEP*-clade, appear to be present in gymnosperms (Carlsbecker *et al.*, 2004). If this phylogenetic interpretation is correct, the *SEP*-clade would seem to have been lost from the extant gymnosperm lineage, rather than generated by a gene duplication event in the angiosperm lineage. In this case, the origin of the angiosperm *SEP*-clade could not have correlated with that of the flower. Even if the above phylogenetic interpretation is incorrect, and *SEP* genes did originate specifically in the angiosperm lineage, the presence of two distinct *SEP*-lineages in ANA grade angiosperms suggests the *SEP*-clade to have existed for some time prior to the origin of the flower. If indeed the *SEP*-clade considerably predates the flower, the formation of quaternary MADS-box complexes involving SEP proteins may also be far more ancient than the flower itself, and hence perhaps not a key factor in flower origin. Furthermore, the formation of quaternary MADS-box complexes may not involve SEP proteins in all plant groups. For example, AGL6 proteins, as close relatives of SEP proteins, might participate in quaternary MADS-box complex formation in gymnosperms. Careful attention should now be paid to the formation of complexes of MADS-box proteins in both angiosperms and gymnosperms

to attempt to address these questions. At present, it is not clear whether the origin of quaternary MADS-box protein complexes correlated with the origin of the flower.

1.7 Carpel specification in monocots

The monocots form a large, monophyletic group of angiosperms whose lineage diverged after the divergence of the ANA grade lineages, probably around 145 MYA (Davies *et al.*, 2004). The functional comparison of carpel development genes in monocot and eudicot models provides evidence of differences in the molecular mechanisms specifying carpel development in these two groups. Genes controlling floral organ identity have been analyzed at a functional level in rice and maize of the Poaceae or grass family. Phylogenetic analyses suggest at least one major gene duplication event to have occurred in the MADS-box C-clade prior to the separation of the rice and maize lineages, with an additional duplication in one of the two sub-clades generated in that duplication, specifically in the maize lineage. Accordingly, the rice C-clade gene *OsMADS58* appears orthologous to the maize gene *ZAG1*, while *OsMADS3* from rice appears orthologous to both *ZMM2* and *ZMM23* from maize (Yamaguchi *et al.*, 2006).

The phenotypes associated with mutations in C-clade genes have been investigated in both rice and maize, though more thoroughly in the former of these species. The inactivation of *OsMADS58* in rice leads to defects in carpel development, though it does not eliminate carpels (Yamaguchi *et al.*, 2006) (Fig. 1.4d). In addition to abnormal carpels, *osmads58* mutants show reduced floral determinacy, indicating a major contribution of this gene to the so-called 'stop' function, which arrests the proliferation of organs in the fourth whorl. Whereas *OSMADS58* appear to act mainly in the fourth floral whorl, the inactivation of its paralogue *OsMADS3* has little or no effect on either carpel development or floral determinacy. Instead, stamen development is eliminated in *osmads3* mutants (Kang *et al.*, 1998; Yamaguchi *et al.*, 2006). Rice plants in which both *OsMADS3* and *OsMADS58* have been inactivated produce aberrant carpels, similar to those of the *osmads58* mutant, suggesting *OsMADS3* to make no unique contribution to carpel development (Yamaguchi *et al.*, 2006). In maize, *zag1* mutants show a defect in floral determinacy, indicating functional conservation of *ZAG1* with its rice orthologue *OsMADS58.* It seems, therefore, that C-clade genes in the grass family have undergone significant sub-functionalization, following a monocot-specific gene duplication. The multiple roles of the single *Arabidopsis* gene *AG* in carpel development, stamen development and floral determinacy are thus shared in a whorl-specific manner between two and three C-clade genes in rice and maize, respectively. Additionally, in these monocot species, carpel development can occur independently of C-clade MADS-box genes, suggesting that some other factor may be involved in the specification of carpel development.

The functions of a paralogous pair of D-clade genes, *OsMADS13* and *OsMADS21,* have been investigated in rice (Dreni *et al.*, 2007). Of these,

OsMADS13 shows ovule-specific expression and appears to play a classical D-function role in ovule development: inactivation of *OsMADS13* results in the ectopic conversion of ovules into internal carpelloid organs. Interestingly, *OsMADS21* appears to make no significant contribution to the D-function, and is expressed more widely in female reproductive tissues. It is thus tempting to speculate that *OsMADS21* might show genetic redundancy in the control of carpel development with rice C-clade MADS-box genes. Double and triple knockout mutants between *osmads21, osmads58* and *osmadsS3* would be required to address this question.

Though the inactivation of C-clade genes does not lead to the loss of carpels in rice flowers, carpels are entirely replaced by ectopic stamens in mutants in which the YABBY gene *DROOPING LEAF (DL)* has been inactivated (Yamaguchi *et al.*, 2004) (Fig. 1.4d). *dl* mutants also show a developmental defect in leaves, which consequently lack a mid-rib. *DL* is the likely rice paralogue of *CRC*, though it clearly shows a considerable functional difference from *CRC* in *Arabidopsis. DL* expression is maintained in the carpels of rice plants in which both *OsMADS3* and *OsMAD58* have been inactivated (Yamaguchi *et al.*, 2006), demonstrating its action to be independent of these C-function genes. This may represent a difference from the situation in *Arabidopsis*, in which *CRC* is a direct target of *AG* (Gomez-Mena *et al.*, 2005). It is not yet clear whether carpel development depends on *DL* expression *per se*, or whether *DL* is mainly responsible for preventing B-function gene expression in the fourth floral whorl. Experiments that combine B-clade, C-clade and *dl* mutations in rice might help to evaluate the role of *DL* in floral patterning, and/or in carpel specification.

The conservation of expression patterns of *CRC* orthologues between *Arabidopsis* and ANA grade angiosperms (Fourquin *et al.*, 2005), as discussed above, suggests the distinct roles of *DL* in carpel identity and leaf development (Yamaguchi *et al.*, 2004) to have arisen specifically in the monocot lineage. Experiments in which *Arabidopsis crc-1* mutants were rescued by transformation with *CRC* orthologues from various species (Fourquin *et al.*, 2007), however, showed the *DL* coding sequence to be capable of restoring near wild-type carpel development when expressed from the *Arabidopsis CRC* promoter. These experiments indicate the *CRC* and *DL* coding sequences to show similar activities in carpel development, suggesting that upstream factors may be largely responsible for the novel function shown by *DL* in carpel specification, or in the definition of the inner limit of the third whorl, in rice. It is not yet clear whether changes to the *DL* coding sequence, in addition to changes to its regulatory region, may have been necessary for the evolution of the leaf development function of *DL* in rice. Transformation of *dl* mutants with constructions containing eudicot *CRC* coding sequences may help to answer this question.

SEP genes, which are necessary for carpel development in eudicots, are also known from monocots. *OsMADS1* from rice corresponds to the *LEAFY HULL STERILE1* locus, and groups within the same clade as *SEP1, SEP2* and *SEP4* from *Arabidopsis* (Zahn *et al.*, 2005). Outer whorl floral organs in

osmads1 loss-of-function mutants take on a leaf-like appearance, whereas the inner whorl floral organs of these mutants are partially converted to paleas and lemmas, which are normally found only in the outer two whorls of rice flowers (Agrawal *et al.*, 2005). These results suggest *OsMADS1* to be a principal component of the E-function in rice (Fig. 1.4d), while the functions of the four remaining rice *SEP* genes, *OsMADS5, OsMADS7, OsMADS8* and *RMADS217* (Zahn *et al.*, 2005), remain to be determined.

In general, carpel and ovule development in the highly derived Poaceae of the monocots seem to depend on the orthologues of regulatory genes that are known to play key roles in these processes in *Arabidopsis* and other eudicots. However, duplications have taken place in several MADS-box gene lineages in Poaceae, including the C-, D- and E-function lineages, in some cases leading to sub-functionalization events between paralogous genes. The precise limits of this sub-functionalization have not yet been defined, which might explain the currently hidden component of the specification of carpel identity in monocots.

1.8 Gene duplication and carpel evolution in the core eudicots

The core eudicots form a monophyletic group that contains several very successful molecular genetic model species, including *Arabidopsis thaliana, Antirrhinum majus* and *Petunia hybrida*. This group is estimated to descend from a last common ancestor that lived around 110 MYA (Davies *et al.*, 2004). Analysis of the *Arabidopsis* genome sequence has provided evidence of a large-scale duplication event that probably occurred not long before the divergence of the main core eudicot lineages, the rosids, asterids and Caryophyllales (De Bodt *et al.*, 2005). As discussed above, several classes of MADS-box genes in core eudicots contain pairs of paralogues that are orthologous to single genes in basal eudicots, such as Ranunculales, and which may therefore have been generated in the hypothesized core eudicot genome duplication event. The retention of many of these pairs of paralogues over long periods of evolutionary time would appear to be a clear indication that sub- and neo-functionalization processes have occurred, rendering both copies of each pair essential or advantageous to survival.

In the core eudicots, two C-clades are present in place of an ancestral *paleoAG* clade in basally diverging eudicot lineages. In *Arabidopsis*, the *euAG* clade contains the *AG* gene itself, while the *PLENA (PLE)* clade (Fig. 1.5), contains a pair of paralogous genes termed *SHATTERPROOF1* and *SHATTERPROOF2 (SHP1/2),* which resulted from a more recent duplication in the *Arabidopsis* lineage. In *Antirrhinum majus*, the probable orthologue of *AG* is termed *FARINELLI (FAR),* while that of *SHP1/2* is the clade-defining gene *PLE.* Interestingly, the non-orthologous genes *AG* and *PLE* are responsible

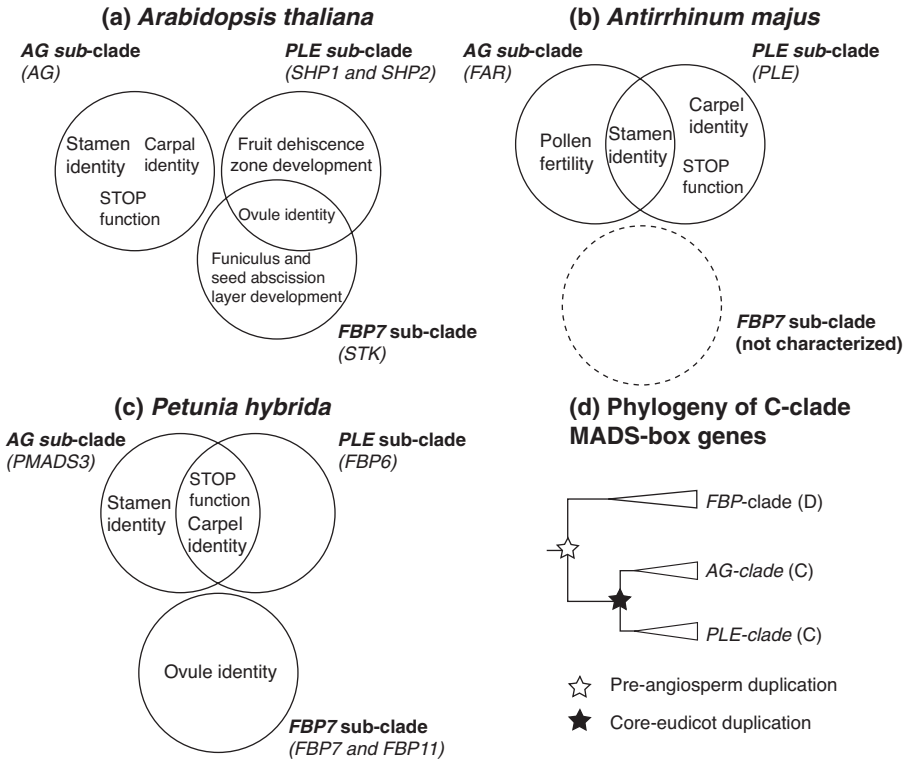

Figure 1.5 Fluidity in the functionalization of C- and D-function MADS-box genes in the core eudicots. (a–c) Venn diagrams representing the functions of genes from the MADS-box clades *AG* (C-function), *PLE* (C-function) and *FBP7* (D-function) in three species of core eudicots. Overlapping regions represent functional redundancy between genes in wild-type genetic backgrounds. *AG*, *AGAMOUS*; *FAR*, *FARINELLI; FBP*, *FLORAL BINDING PROTEIN*; *PLE*, *PLENA*; *SHP*, *SHATTERPROOF*. (d) The sequence of duplications that generated of the eudicot *AG*, *PLE* and *FBP7* MADS-box gene clades.

for specifying the C-function in *Arabidopsis* and *Antirrhinum,* respectively (Davies *et al.*, 1999; Kramer *et al.*, 2004) (Fig. 1.5). *FAR,* by contrast, is redundantly involved in stamen development and is required for pollen fertility in *Antirrhinum.* In an example of neo-functionalization, the paralogues *SHP1* and *SHP2* play a novel role in *Arabidopsis* fruit development (Liljegren *et al.*, 2000). In *Petunia hybrida*, which, as a member of the asterids, is more closely related to *Antirrhinum* than to *Arabidopsis* (Fig. 1.2), a further case of sub-functionalization is apparent. The *Petunia AG* orthologue, *PMADS3,* is principally responsible for stamen development (Kapoor *et al.*, 2002), though it also plays a redundant role with the *PLE* orthologue *FLORAL BINDING PROTEIN6 (FBP6)* in carpel development and floral determinacy (Kramer *et al.*, 2004).

Though sub-functionalization between the paralogous *AG* and *PLE* clades in *Arabidopsis* has left *AG* playing the major C-function role, elegant experiments involving multiple mutants show that the *SHP* genes have retained a capacity for C-function activity, perhaps reflecting the C-function role of the common ancestor, prior to the radiation of the core eudicots, which these genes share with *AG.* Ectopic carpelloid organs develop in the first floral whorl of *Arabidopsis ag* mutants, conditionally on the inactivation of *APETALLA2 (AP2)*, a gene which contributes to the A-function (Bowman *et al.*, 1991). This effect is thought to occur because *AP2* is responsible for down-regulating C-clade MADS-box genes in the outer floral whorls of wild-type plants. In the case of *ag/ap2* double mutants, the C-function activity responsible for specifying ectopic carpel development in the first whorl is provided by *SHP1* and *SHP2,* as evidenced by the fact that first whorl organs of quadruple *ap2/ag/shp1/shp2* mutants are devoid of all carpelloid features (Pinyopich *et al.*, 2003). These data indicate a subtle effect of functional overlap between paralogous gene clades, which does not equate to simple genetic redundancy: the *SHP* genes adopt novel C-function-like roles in *ap2/ag* mutants which they do not play in wild-type plants.

The fluidity of functions among duplicated genes in the core eudicots is further illustrated by an exchange of function between C- and D-clade MADS-box genes. Two paralogous D-function genes in *Petunia, FBP7* and *FBP11*, are redundantly essential for ovule development (Angenent *et al.*, 1995). The probable *Arabidopsis* orthologue of these genes, *SEEDSTICK (STK),* is also involved in ovule development, though *STK* shares this role redundantly with the C-clade genes *SHP1* and *SHP2* (Fig. 1.5). Accordingly, the Arabidopsis *stk/shp1/shp2* triple mutant (Pinyopich *et al.*, 2003), like the *Petunia fpb7/fpb11* double mutant (Angenent *et al.*, 1995), produces supernumerary carpels in the place of ovules within the gynoecium. In addition to its redundant role in ovule specification, *STK* plays non-redundant roles in the development of the funiculus and in seed abscission in *Arabidopsis* (Pinyopich *et al.*, 2003). The combined C+D-clade in the eudicots, whose members were derived from duplication events that occurred both before and after the radiation of the angiosperms, therefore represents a complex situation in which diverse evolutionary processes have taken place. These processes include sub-functionalization between paralogous genes, exchanges of function between paralogous genes, exchanges of function between non-paralogous genes, and, finally, neo-functionalization to generate novel fruit shattering mechanisms (Fig. 1.5).

1.9 The A-function finds a role in fruit development

A further likely consequence of the hypothesized genome duplication at the base of the core eudicots is the generation of a second sub-clade of MADS-box genes within the A-clade (Litt and Irish, 2003). The A-function MADS-box gene *APETALLA1 (AP1)* plays roles in floral meristem patterning and in the

specification of perianth (petal and sepal) organ identity in *Arabidopsis*. This latter role corresponds to the A-function, as defined by the ABC and ABCE models. However, gene duplications in the core eudicots have provided further A-clade sequences, one of which appears to have been recruited to carpel and fruit development somewhere along the lineage leading to *Arabidopsis*. Accordingly, the *Arabidopsis* A-clade MADS-box gene *FRUITFULL (FUL)* is involved in the patterning of the gynoecium and fruit wall (Gu *et al.*, 1998). *FUL* is known to act in a network involving a large number of genes (Roeder *et al.*, 2003; Liljegren *et al.*, 2004), including the MADS-box genes *SHP1* and *SHP2* (Ferrandiz *et al.*, 2000) that also function redundantly with *STK* in ovule development, as described above. Gene duplication in the A-function clade of MADS-box genes has thus resulted in novel fruit shattering mechanisms in the Brassicaceae by the process of neo-functionalization.

An interesting feature of gene-duplication in the A-clade is the evolution of a distinct C-terminal protein motif in *AP1* genes, apparently produced by a frame-shift mutation that occurred towards the 3′-extremity of the coding sequence in an ancestor of the core eudicots (Litt and Irish, 2003). This frame-shift created a farnesylsation site in the encoded protein that is known to be post-translationally modified in vivo in *Arabidopsis* and which is required for wild-type *AP1* protein activity (Yalovsky *et al.*, 2000). Other frame-shift mutations in duplicated genes are present in the B- and C-function MADS-box clades of the eudicots (Vandenbussche *et al.*, 2003). However, the conserved motifs generated in these cases are distinct from that of the *AP1* lineage and do not contain farnesylation sites. The novel C-terminal motifs present in certain lineages within the eudicot A-, B- and C-clades of MADS-box genes have been conserved over a long period, clearly indicating their functional significance. However, it is not yet known whether the functions of these novel motifs are connected with biochemical processes in common, such as the higher order assembly or sequestration of MADS-box transcription factor complexes (Vandenbussche *et al.*, 2003).

1.10 The multiple origins and mechanisms of syncarpy in the angiosperms

Most angiosperm flowers possess more than one carpel. As discussed above, the carpels of species from the early-diverging ANA grade lineages are typically separate structures that occur in a spiral arrangement at the centre of the flower (Fig. 1.3a–f). Such an arrangement, with separate carpels, is termed apocarpic and, from its presence in early-diverging lineages, appears to represent the pleisiomorphic condition of the angiosperms. However, more than 80% of extant angiosperm species are syncarpic: their carpels are fused into a single female structure in the centre of the flower (Endress, 1982) (Fig. 1.3g–h).

Various morphological sequences have been described which lead to carpel fusion in syncarpic species (Verbeke, 1992). These developmental processes

can be conceptually divided into two types, based on the timing of the fusion event involved. Accordingly, in cases in which carpels are fused from the earliest emergence of carpel primordia, the organ fusion is termed 'congenital', whereas in cases in which carpels are observed to fuse together during development, the fusion is termed 'post-genital'. Congenital carpel fusion is the most common type, whereas post-genital carpel fusion is known from only a handful of angiosperm families (Lolle and Pruitt, 1999). *Arabidopsis,* which typifies gynoecium development in the Brassicaceae, is a good example of congenital carpel fusion. At developmental stage 6 of *Arabidopsis* flowers (Smyth *et al.*, 1990), a central dome emerges from the flower meristem which will become the syncarpic gynoecium. By developmental stage 7, this dome begins to invaginate by a reduction in growth rate at its centre to generate a slot, which forms in line with the pair of lateral stamen primordia of the *Arabidopsis* flower bud. The tissues on either side of this slot, perpendicular to the lateral stamen primordia, then grow out to meet in the middle, eventually forming a vertical septum that divides the ovary of the *Arabidopsis* gynoecium into two locules. The gynoecium wall undergoes considerable differentiation at the extremities of the vertical septum to define the dehiscence zones that will permit pod shattering in the mature fruit.

Interestingly, though *Arabidopsis* is a clear example of a species showing congenital syncarpy, it is not absolutely clear how many carpels are fused together in the *Arabidopsis* gynoecium. Most recent authors (Okada *et al.*, 1989) have interpreted this structure as containing two carpels, corresponding to the two locules of the (secondarily) divided ovary. This view is to some extent supported by molecular genetic studies, which have succeeded in isolating a number of mutants in which carpel fusion is affected, resulting in the division of the *Arabidopsis* gynoecium into two separate carpelloid organs. However, the ovary wall contains four vascular traces in *Arabidopsis*, two of which occur at either extremity of the vertical septum, and two in positions corresponding to the valves of the ovary. Hence, it is entirely possible that there are four carpels in the wild-type *Arabidopsis* gynoecium (Lawrence, 1951). According to this view, the valve carpels would have become sterile, whereas the carpels at either extremity of the vertical septum, corresponding to the positions of the placentae, would have remained fertile.

Postgenital syncarpy has been best characterized in *Catharanthus roseus* (The Madagascar Periwinkle) of the Apocynaceae. In this species, two separate carpel primordia are initiated and grow until their inner surfaces come into contact (Walker, 1978; Siegel and Verbeke, 1989; Verbeke, 1992). The already differentiated epidermal cells of these surfaces then begin to interlock and redifferentiate into parenchyma. This redifferentiation is dependent on diffusible, water-soluble substances produced by the carpels themselves and takes a total of about 9 h (Siegel and Verbeke, 1989). Even before epidermal cell redifferentiation has terminated, plasmodesmata can be observed to form between the surface layers of the fusing carpels (Vanderschoot *et al.*, 1995). Walker (1978) has shown that the pollen tubes are then able to cross between

the carpels, thus fulfilling one of the principal theoretical reasons for the evolution of syncarpy, which are explained in greater detail below.

The major evolutionary advantage of syncarpy is probably to allow a regular repartition of pollen tubes within the gynoecium by the formation of a 'compitum', which consists of tissues that permit pollen tube transfer between carpels (Endress, 1982). Accordingly, a visit to a syncarpic by a single pollinator might result in full seed set. Another potentially important advantage of syncarpy results from the enhanced competition between pollen tubes that this phenomenon produces, which may act as a filter of fitness by selecting for vigorous male parents. Syncarpy also allows the production of larger fruits, with potentially more complex and efficient seed dispersal mechanisms (Walker, 1978; Endress, 1982; Armbruster *et al.*, 2002). A further advantage of syncarpy may stem from its potentially lesser requirement for cell wall synthesis, compared to an apocarpic gynoecium of similar total size.

Though most representatives of basally diverging angiosperm lineages are morphologically apocarpous, Endress (1982) has noted several unusual constructions that confer a degree of functional syncarpy. In *Nymphaea* (Nymphaeaceae), for example, large quantities of mucilage are secreted from the stigmatic tissues of the separate carpels. Pollen tubes are able to grow through the resulting mucilage layer, which is sufficiently extensive as to form a bridge between the carpels. In *Tambourissa* of the eumagnolid family Monimiaceae (Laurales), a 'hyperstigma' is generated by the entire female flower, which forms a cup that fills with mulicage into which pollen grains may fall and germinate (Endress, 1982). The separate carpels of *Tambourissa* develop in this floral cup, and are thus potentially accessible to any pollen grain present. Functional syncarpy in the ANA grade genus *Illicium* (Austrobaileyales) is achieved by the growth of pollen tubes in a groove around a central axis which connects the separate carpels (Williams *et al.*, 1993). By mapping different arrangements of morphological and functional syncarpy onto a phylogeny of the angiosperms, Armbruster *et al.* (2002) estimated that there have been at least 17 independent transitions from apocarpy to syncarpy during angiosperm evolution. Only two instances, by contrast, of a likely change from syncarpy to morphological apocarpy were noted.

As syncarpy seems to have arisen several times independently in the angiosperms, it is possible that distinct molecular mechanisms have been recruited to bring about carpel fusion in distantly related syncarpic groups. In *Arabidopsis*, numerous mutations, including *aintegumenta, crabs claw, ettin, leunig, spatula* and *tousled*, generate various degrees of carpel separation, either singly or in double mutant combinations (Sessions and Zambryski, 1995; Roe *et al.*, 1997; Alvarez and Smyth, 1999; Liu *et al.*, 2000). It is not, however, clear whether these genes played any role in the evolution of syncarpy in the *Arabidopsis* lineage. Rather, the inactivation of such key regulators of carpel development might disrupt the delicate and labile process of carpel fusion, which perhaps evolved through other, unrelated molecular changes. One genetic system that may potentially have been responsible for the generation of

congenital syncarpy, however, does seem worthy of particular mention. The genes *CUP-SHAPED COTYLEDON1* and *2 (CUC1/2)* (Aida *et al.*, 1997) are members of the *NAC* transcription factor family which promote, in a partially redundant manner, organ separation in a number of positions in the plant, including the cotyledons, floral meristems and leaf margins. Both *CUC1* and *CUC2* are, together with five other *NAC* genes, negatively regulated by the *MIR164* family of microRNAs, which contains three genes in *Arabidopsis.* Interestingly, both the triple mutant *mir164abc* (Sieber *et al.*, 2007), and plants transformed with a *miR164*-resistant allele of *CUC2* (Nikovics *et al.*, 2006), show complete carpel separation, thus implicating the negative regulation of *CUC2* by *miR164* in the process of carpel fusion in *Arabidopsis.* The *CUC* and *MIR164* system may represent a conserved developmental module that has been recruited many times independently over the course of angiosperm evolution to modify certain highly variable traits in the angiosperms, such as leaf dissection (Nikovics *et al.*, 2006). It is therefore possible that this developmental module has also been independently recruited to generate syncarpy in distinct angiosperm lineages. The transformation of diverse syncarpic species with *miR164*-resistant alleles of their own native *CUC* genes could help to shed some light on this question.

A clue to the potential molecular mechanism of post-genital carpel fusion in *C. roseus* and other taxa comes from the *Arabidopsis* mutant *fiddlehead (fdh)*. In *fdh* mutants, all above-ground organ types, including leaves, stems, sepals, petals and stamens, tend to fuse together on contact (Lolle *et al.*, 1992). In addition, wild-type *Arabidopsis* pollen can readily germinate and emit pollen tubes into leaf and other non-carpel tissues of *fdh* mutants (Lolle and Cheung, 1993). *FDH* encodes an enzyme necessary for the generation of the waxy cuticle (Pruitt *et al.*, 2000). Inactivation of *FDH* makes the cuticle much more permeable to small molecules, and this effect also seems to be conserved in *Antirrhinum majus* (Efremova *et al.*, 2004). Interestingly, the transfer of small water-soluble molecules is known to be involved in post-genital carpel fusion in *C. roseus* (Siegel and Verbeke, 1989). It thus seems plausible that carpel fusion in *C. roseus* and other taxa showing post-genital fusion might depend on the down-regulation, specifically in the contacting surface of developing carpels, of a gene such as *FDH* that is necessary for cuticle formation.

1.11 A fruit by any other name: evolutionary convergence between angiosperms and gymnosperms

The fruit is a pleisiomorphic character of the angiosperms, and to use this term to describe non-angiosperm seed bearing structures might be regarded as botanical heresy! However, it is interesting to note that the reproductive structures of several groups of living gymnosperms have evolved to superficially resemble angiosperm fruits. For example, the aril of the yew, *Taxus baccata* (Taxaceae), is red and fleshy at maturity, and contributes to

seed dissemination by forming a food source for birds. Arils in Taxaceae and Taxodiaceae develop from ovuliferous scales that grow around the seed after pollination. These structures are thus quite distinct from angiosperm fruits, which are formed by the modification of pre-existing carpel tissues. Similarly, in *Gnetum* of the gymnosperm order Gnetales, a fruit-like structure is generated from two extra tissue layers that form around the unitegmetic ovule. In the case of Gnetales, the production of fruit-like structures represents just one of a long list of evolutionary convergences with angiosperms (Donoghue and Doyle, 2000). For example, a form of double fertilization is present in this group, though this leads to the production of a second, inviable embryo, rather than endosperm tissue (Friedman, 1998). Other evolutionary convergences of Gnetales with angiosperms include an apical meristem that is divided into a tunica and corpus, vessel elements in xylem tissues, a lack of archegonia, net-veined leaves (in *Gnetum*), morphologically bisexual male strobili (in *Gnetum* and *Welwitschia*), and even insect pollination (in some *Gnetum* spp.) (Kato *et al.*, 1995). Structures playing the role of fruits, but which are not derived from carpels, have also evolved in some angiosperm groups. Examples of such false fruits include the 'pome', which develops from the receptacle in Rosaceae including *Malus and Pyrus*, and the false berry of *Vaccinium* spp.

References

Agrawal, G., Abe, K., Yamazaki, M., Miyao, A. and Hirochika, H. (2005) Conservation of the E-function for floral organ identity in rice revealed by the analysis of tissue culture-induced loss-of-function mutants of the OsMADS1 gene. *Plant Molecular Biology* **59**, 125–135.

Aida, M., Ishida, T., Fukaki, H., Fujisawa, H. and Tasaka, M. (1997) Genes involved in organ separation in Arabidopsis: an analysis of the cup-shaped cotyledon mutant. *Plant Cell* **9**, 841–857.

Albert, V.A., Soltis, D.E., Carlson, J.E., Farmerie, W.G., Wall, P.K., Ilut, D.C., Solow, T.M., Mueller, L.A., Landherr, L.L., Hu, Y., Buzgo, M., Kim, S., Yoo, M.J., Frohlich, M.W., Perl-Treves, R., Schlarbaum, S.E., Bliss, B.J., Zhang, X., Tanksley, S.D., Oppenheimer, D.G., Soltis, P.S., Ma, H., DePamphilis, C.W. and Leebens-Mack, J.H. (2005) Floral gene resources from basal angiosperms for comparative genomics research. *BMC Plant Biology* **5**, 5.

Alvarez, J. and Smyth, D.R. (1999) CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS. *Development* **126**, 2377–2386.

Angenent, G.C., Franken, J., Busscher, M., Vandijken, A., Vanwent, J.L., Dons, H.J.M. and Vantunen, A.J. (1995) A novel class of MADS box genes is involved in ovule development in petunia. *Plant Cell* **7**, 1569–1582.

Aoki, S., Uehara, K., Imafuku, M., Hasebe, M. and Ito, M. (2004) Phylogeny and divergence of basal angiosperms inferred from APETALA3- and PISTILLATA-like MADS-box genes. *Journal of Plant Research* **117**, 229–244.

Arber, E.A.N. and Parkin, J. (1907) On the origin of angiosperms. *Botanical Journal of the Linnean Society* **38**, 29–80.

Armbruster, W.S., Debevec, E.M. and Willson, M.F. (2002) Evolution of syncarpy in angiosperms: theoretical and phylogenetic analyses of the effects of carpel fusion on offspring quantity and quality. *Journal of Evolutionary Biology* **15**, 657–672.

Barrett, P.M. and Willis, K.J. (2001) Did dinosaurs invent flowers? Dinosaur–angiosperm coevolution revisited. *Biological Reviews* **76**, 411–447.

Baum, D.A. and Hileman, L.C. (2006) A developmental genetic model for the origin of the flower. In: *Flowering and Its Manipulation* (ed. C. Ainsworth). Blackwell, Oxford.

Becker, A. and 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.

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

Bennett, M.D., Leitch, I.J., Price, H.J. and Johnston, J.S. (2003) Comparisons with Caenorhabditis (similar to 100 Mb) and Drosophila (similar to 175 Mb) using flow cytometry show genome size in Arabidopsis to be similar to 157 Mb and thus similar to 25% larger than the Arabidopsis genome initiative estimate of similar to 125 Mb. *Annals of Botany* **91**, 547–557.

Bernhardt, P., Sage, T., Weston, P., Azuma, H., Lam, M., Thien, L.B. and Bruhl, J. (2003) The pollination of Trimenia moorei (Trimeniaceae): floral volatiles, insect/wind pollen vectors and stigmatic self-incompatibility in a basal angiosperm. *Annals of Botany* **92**, 445–458.

Bowman, J.L. (2000) The YABBY gene family and abaxial cell fate. *Current Opinion in Plant Biology* **3**, 17–22.

Bowman, J.L. and Smyth, D.R. (1999) CRABS CLAW, a gene that regulates carpel and nectary development in Arabidopsis, encodes a novel protein with zinc finger and helix–loop–helix domains. *Development* **126**, 2387–2396.

Bowman, J.L., Smyth, D.R. and Meyerowitz, E.M. (1991) Genetic interactions among floral homeotic genes of arabidopsis. *Development* **112**, 1–20.

Brenner, E.D., Katari, M.S., Stevenson, D.W., Rudd, S.A., Douglas, A.W., Moss, W.N., Twigg, R.W., Runko, S.J., Stellari, G.M., McCombie, W.R. and Coruzzi, G.M. (2005) EST analysis in Ginkgo biloba: an assessment of conserved developmental regulators and gymnosperm specific genes. *BMC Genomics* **6**, 143.

Brenner, E.D., Stevenson, D.W., McCombie, R.W., Katari, M.S., Rudd, S.A., Mayer, K.F.X., Palenchar, P.M., Runko, S.J., Twigg, R.W., Dai, G.W., Martienssen, R.A., Benfey, P.N. and Coruzzi, G.M. (2003) Expressed sequence tag analysis in Cycas, the most primitive living seed plant. *Genome Biology* **4**, R78.

Carlsbecker, A., Tandre, K., Johanson, U., Englund, M. and Engstrom, P. (2004) The MADS-box gene DAL1 is a potential mediator of the juvenile-to-adult transition in Norway spruce (Picea abies). *Plant Journal* **40**, 546–557.

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

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

Davies, B., Motte, P., Keck, E., Saedler, H., Sommer, H. and Schwarz-Sommer, Z. (1999) PLENA and FARINELLI: redundancy and regulatory interactions between two Antirrhinum MADS-box factors controlling flower development. *Embo Journal* **18**, 4023–4034.

Davies, T.J., Barraclough, T.G., Chase, M.W., Soltis, P.S., Soltis, D.E. and Savolainen, V. (2004) Darwin's abominable mystery: insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 1904–1909.

De Bodt, S., Maere, S. and Van de Peer, Y. (2005) Genome duplication and the origin of angiosperms. *Trends in Ecology and Evolution* **20**, 591–597.

Donoghue, M.J. and Doyle, J.A. (2000) Seed plant phylogeny: demise of the anthophyte hypothesis? *Current Biology* **10**, R106–R109.

Dornelas, M.C. and Rodriguez, A.P.M. (2005) A FLORICAULA/LEAFY gene homolog is preferentially expressed in developing female cones of the tropical pine Pinus caribaea var. caribaea. *Genetics and Molecular Biology* **28**, 299–307.

Dreni, L., Jacchia, S., Fornara, F., Fornari, M., Ouwerkerk, P.B.F., An, G.H., Colombo, L. and Kater, M.M. (2007) The D-lineage MADS-box gene OsMADS13 controls ovule identity in rice. *Plant Journal* **52**, 690–699.

Efremova, N., Schreiber, L., Bar, S., Heidmann, I., Huijser, P., Wellesen, K., Schwarz-Sommer, Z., Saedler, H. and Yephremov, A. (2004) Functional conservation and maintenance of expression pattern of FIDDLEHEAD-like genes in Arabidopsis and Antirrhinum. *Plant Molecular Biology* **56**, 821–837.

Endress, P.K. (1982) Syncarpy and alternative modes of escaping disadvantages of apocarpy in primitive angiosperms. *Taxon* **31**, 48–52.

Endress, P.K. (2001) The flowers in extant basal angiosperms and inferences on ancestral flowers. *International Journal of Plant Sciences* **162**, 1111–1140.

Endress, P.K. and Igersheim, A. (2000) Gynoecium structure and evolution in basal angiosperms. *International Journal of Plant Sciences* **161**, S211–S223.

Fahn, A. (1982) *Plant Anatomy*. Pergamon, Oxford.

Feild, T.S., Arens, N.C. and Dawson, T.E. (2003) The ancestral ecology of angiosperms: emerging perspectives from extant basal lineages. *International Journal of Plant Sciences* **164**, S129–S142.

Feild, T.S., Arens, N.C., Doyle, J.A., Dawson, T.E. and Donoghue, M.J. (2004) Dark and disturbed: a new image of early angiosperm ecology. *Paleobiology* **30**, 82–107.

Ferrandiz, C., Liljegren, S.J. and Yanofsky, M.F. (2000) Negative regulation of the SHATTERPROOF genes by FRUITFULL during Arabidopsis fruit development. *Science* **289**, 436–438.

Fourquin, C., Vinauger-Douard, M., Chambrier, P., Berne-Dedieu, A. and Scutt, C.P. (2007) Functional conservation between CRABS CLAW orthologues from widely diverged Angiosperms. *Annals of Botany* **100**, 651–657.

Fourquin, C., Vinauger-Douard, M., Fogliani, B., Dumas, C. and Scutt, C.P. (2005) Evidence that CRABS CLAW and TOUSLED have conserved their roles in carpel development since the ancestor of the extant angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 4649–4654.

Friedman, W.E. (1998) The evolution of double fertilization and endosperm: an 'historical' perspective. *Sexual Plant Reproduction* **11**, 6–16.

Friedman, W.E. (2006) Embryological evidence for developmental lability during early angiosperm evolution. *Nature* **441**, 337–340.

Friedman, W.E. (2008) Hydatellaceae are water lilies with gymnospermous tendencies. *Nature* **453**, 94–97.

Friis, E.M., Pedersen, K.R. and Crane, P.R. (2005) When Earth started blooming: insights from the fossil record. *Current Opinion in Plant Biology* **8**, 5–12.

Frohlich, M.W. (2003) An evolutionary scenario for the origin of flowers. *Nature Reviews Genetics* **4**, 559–566.

Frohlich, M.W. and Parker, D.S. (2000) The mostly male theory of flower evolutionary origins: from genes to fossils. *Systematic Botany* **25**, 155–170.

Gomez-Mena, C., de Folter, S., Costa, M.M.R., Angenent, G.C. and Sablowski, R. (2005) Transcriptional program controlled by the floral homeotic gene AGAMOUS during early organogenesis. *Development* **132**, 429–438.

Goremykin, V., Bobrova, V., Pahnke, J., Troitsky, A., Antonov, A. and Martin, W. (1996) Noncoding sequences from the slowly evolving chloroplast inverted repeat in addition to rbcL data do not support Gnetalean affinities of angiosperms. *Molecular Biology and Evolution* **13**, 383–396.

Goremykin, V.V., Hansmann, S. and Martin, W.F. (1997) Evolutionary analysis of 58 proteins encoded in six completely sequenced chloroplast genomes: revised molecular estimates of two seed plant divergence times. *Plant Systematics and Evolution* **206**, 337–351.

Grotkopp, E., Rejmanek, M., Sanderson, M.J. and Rost, T.L. (2004) Evolution of genome size in pines (Pinus) and its life-history correlates: supertree analyses. *Evolution* **58**, 1705–1729.

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

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

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

Kang, H.G., Jeon, J.S., Lee, S. and An, G.H. (1998) Identification of class B and class C floral organ identity genes from rice plants. *Plant Molecular Biology* **38**, 1021–1029.

Kapoor, M., Tsuda, S., Tanaka, Y., Mayama, T., Okuyama, Y., Tsuchimoto, S. and Takatsuji, H. (2002) Role of petunia pMADS3 in determination of floral organ and meristem identity, as revealed by its loss of function. *Plant Journal* **32**, 115–127.

Kato, M., Inoue, T. and Nagamitsu, T. (1995) Pollination biology of gnetum (Gnetaceae) in a lowland mixed dipterocarp forest in sarawak. *American Journal of Botany* **82**, 862–868.

Kim, S., Koh, J., Ma, H., Hu, Y., Endress, P.K., Hauser, B.A., Buzgo, M., Soltis, P.S. and Soltis, D.E. (2005) Sequence and expression studies of A-, B-, and E-class MADS-box homologues in Eupomatia (Eupomatiaceae): support for the bracteate origin of the calyptra. *International Journal of Plant Sciences* **166**, 185–198.

Kramer, E.M. and Hall, J.C. (2005) Evolutionary dynamics of genes controlling floral development. *Current Opinion in Plant Biology* **8**, 13–18.

Kramer, E.M., Jaramillo, M.A. and Di Stilio, V.S. (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.

Kuzoff, R.K. and Gasser, C.S. (2000) Recent progress in reconstructing angiosperm phylogeny. *Trends in Plant Science* **5**, 330–336.

Lawrence, G.H. (1951) *Taxonomy of Flowering Plants*. Macmillan, New York.

Leitch, I.J. and Hanson, L. (2002) DNA C-values in seven families fill phylogenetic gaps in the basal angiosperms. *Botanical Journal of the Linnean Society* **140**, 175–179.

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

Liljegren, S.J., Roeder, A.H.K., Kempin, S.A., Gremski, K., Ostergaard, L., Guimil, S., Reyes, D.K. and Yanofsky, M.F. (2004) Control of fruit patterning in Arabidopsis by INDEHISCENT. *Cell* **116**, 843–853.

Litt, A. and Irish, V.F. (2003) Duplication and diversification in the APETALA1/FRUITFULL floral homeotic gene lineage: implications for the evolution of floral development. *Genetics* **165**, 821–833.

Liu, Z.C., Franks, R.G. and Klink, V.P. (2000) Regulation of gynoecium marginal tissue formation by LEUNIG and AINTEGUMENTA. *Plant Cell* **12**, 1879–1891.

Lolle, S.J. and Cheung, A.Y. (1993) Promiscuous germination and growth of wildtype pollen from Arabidopsis and related species on the shoot of the Arabidopsis Mutant, Fiddlehead. *Developmental Biology* **155**, 250–258.

Lolle, S.J., Cheung, A.Y. and Sussex, I.M. (1992) Fiddlehead – an Arabidopsis mutant constitutively expressing an organ fusion program that involves interactions between epidermal-cells. *Developmental Biology* **152**, 383–392.

Lolle, S.J. and Pruitt, R.E. (1999) Epidermal cell interactions: a case for local talk. *Trends in Plant Science* **4**, 14–20.

Moore, M.J., Bell, C.D., Soltis, P.S. and Soltis, D.E. (2007) Using plastid genome-scale data to resolve enigmatic relationships among basal angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 19363–19368.

Mouradov, A., Glassick, T., Hamdorf, B., Murphy, L., Fowler, B., Maria, S. and Teasdale, R.D. (1998) NEEDLY, a Pinus radiata ortholog of FLORICAULA/LEAFY genes, expressed in both reproductive and vegetative meristems. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 6537–6542.

Nesi, N., Debeaujon, I., Jond, C., Stewart, A.J., Jenkins, G.I., Caboche, M. and Lepiniec, L. (2002) The TRANSPARENT TESTA16 locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. *Plant Cell* **14**, 2463–2479.

Nikovics, K., Blein, T., Peaucelle, A., Ishida, T., Morin, H., Aida, M. and Laufs, P. (2006) The balance between the MIR164A and CUC2 genes controls leaf margin serration in Arabidopsis. *Plant Cell* **18**, 2929–2945.

Ohri, D. and Khoshoo, T.N. (1986) Genome size in gymnosperms. *Plant Systematics and Evolution* **153**, 119–132.

Okada, K., Komaki, M.K. and Shimura, Y. (1989) Mutational analysis of pistil structure and development of Arabidopsis-Thaliana. *Cell Differentiation and Development* **28**, 27–38.

Page, C.N. (1990) Coniferophytina. In: *Pteridophytes and Gymnosperms* (eds K.U. Kramer, and P.S. Green). Springer, Berlin.

Pavy, N., Johnson, J.J., Crow, J.A., Paule, C., Kunau, T., MacKay, J. and Retzel, E.F. (2007) ForestTreeDB: a database dedicated to the mining of tree transcriptomes. *Nucleic Acids Research* **35**, D888–D894.

Pavy, N., Paule, C., Parsons, L., Crow, J.A., Morency, M.J., Cooke, J., Johnson, J.E., Noumen, E., Guillet-Claude, C., Butterfield, Y., Barber, S., Yang, G., Liu, J., Stott, J., Kirkpatrick, R., Siddiqui, A., Holt, R., Marra, M., Seguin, A., Retzel, E., Bousquet, J. and MacKay, J. (2005) Generation, annotation, analysis and database integration of 16,500 white spruce EST clusters. *BMC Genomics* **6**, 144.

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

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

Prakash, N. and Alexander, J.H. (1984) Self-incompatibility in *Austrobaileya scandens*. In: *Pollination '84* (eds E.G. Williams and R.B. Knox). University of Melbourne, Melbourne.

Pruitt, R.E., Vielle-Calzada, J.P., Ploense, S.E., Grossniklaus, U. and Lolle, S.J. (2000) FIDDLEHEAD, a gene required to suppress epidermal cell interactions in Arabidopsis, encodes a putative lipid biosynthetic enzyme. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 1311–1316.

Qiu, Y.L., Lee, J., Whitlock, B.A., Bernasconi-Quadroni, F. and Dombrovska, O. (2001) Was the ANITA rooting of the angiosperm phylogeny affected by long-branch attraction? *Molecular Biology and Evolution* **18**, 1745–1753.

Roe, J.L., Nemhauser, J.L. and Zambryski, P.C. (1997) TOUSLED participates in apical tissue formation during gynoecium development in Arabidopsis. *Plant Cell* **9**, 335–353.

Roeder, A.H.K., Ferrandiz, C. and Yanofsky, M.F. (2003) The role of the REPLUMLESS homeodomain protein in patterning the Arabidopsis fruit. *Current Biology* **13**, 1630–1635.

Saarela, J.M., Rai, H.S., Doyle, J.A., Endress, P.K., Mathews, S., Marchant, A.D., Briggs, B.G. and Graham, S.W. (2007) Hydatellaceae identified as a new branch near the base of the angiosperm phylogenetic tree. *Nature* **446**, 312–315.

Savard, L., Li, P., Strauss, S.H., Chase, M.W., Michaud, M. and Bousquet, J. (1994) Chloroplast and nuclear gene-sequences indicate late Pennsylvanian time for the last common ancestor of extant seed plants. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 5163–5167.

Scutt, C.P., Vinauger-Douard, M., Fourquin, C., Ailhas, J., Kuno, N., Uchida, K., Gaude, T., Furuya, M. and Dumas, C. (2003) The identification of candidate genes for a reverse genetic analysis of development and function in the Arabidopsis gynoecium. *Plant Physiology* **132**, 653–665.

Sessions, R.A. and Zambryski, P.C. (1995) Arabidopsis gynoecium structure in the wild-type and in effin mutants. *Development* **121**, 1519–1532.

Sieber, P., Wellmer, F., Gheyselinck, J., Riechmann, J.L. and Meyerowitz, E.M. (2007) Redundancy and specialization among plant microRNAs: role of the MIR164 family in developmental robustness. *Development* **134**, 1051–1060.

Siegel, B.A. and Verbeke, J.A. (1989) Diffusible factors essential for epidermal-cell redifferentiation in catharanthus-roseus. *Science* **244**, 580–582.

Smyth, D.R., Bowman, J.L. and Meyerowitz, E.M. (1990) Early flower development in Arabidopsis. *Plant Cell* **2**, 755–767.

Soltis, D.E., Albert, V.A., Leebens-Mack, J., Palmer, J.D., Wing, R.A., Depamphilis, C.W., Ma, H., Carlson, J.E., Altman, N., Kim, S., Wall, P.K., Zuccolo, A. and Soltis, P.S. (2008) The Amborella genome: an evolutionary reference for plant biology. *Genome Biology* **9**, 402.

Stellari, G.M., Jaramillo, M.A. and Kramer, E.M. (2004) Evolution of the APETALA3 and PISTILLATA lineages of MADS-box-containing genes in the basal angiosperms. *Molecular Biology and Evolution* **21**, 506–519.

Tandre, K., Albert, V.A., Sundas, A. and Engstrom, P. (1995) Conifer homologs to genes that control floral development in angiosperms. *Plant Molecular Biology* **27**, 69–78.

Tandre, K., Svenson, M., Svensson, M.E. and Engstrom, P. (1998) Conservation of gene structure and activity in the regulation of reproductive organ development of conifers and angiosperms. *Plant Journal* **15**, 615–623.

Theissen, G., Becker, A., Winter, K.-U., Munster, T., Kirchner, C. and Saedler, H. (2002) How the land plants learned their floral ABCs: the role of MADS-box genes in the evolutionary origin of flowers. In: *Developmental Genetics and Plant Evolution* (eds Q.B.C. Cronk, R.M. Bateman and J.A. Hawkins). Taylor and Francis, London.

Theissen, G. and Melzer, R. (2007) Molecular mechanisms underlying origin and diversification of the angiosperm flower. *Annals of Botany* **100**, 603–619.

Theissen, G. and Saedler, H. (2001) Plant biology – floral quartets. *Nature* **409**, 469–471.

Vandenbussche, M., Theissen, G., Van de Peer, Y. and Gerats, T. (2003) Structural diversification and neo-functionalization during floral MADS-box gene evolution by C-terminal frameshift mutations. *Nucleic Acids Research* **31**, 4401–4409.

Vanderschoot, C., Dietrich, M.A., Storms, M., Verbeke, J.A. and Lucas, W.J. (1995) Establishment of a cell-to-cell communication pathway between separate carpels during gynoecium development. *Planta* **195**, 450–455.

Vazquez-Lobo, A., Carlsbecker, A., Vergara-Silva, F., Alvarez-Buylla, E.R., Pinero, D. and Engstrom, P. (2007) Characterization of the expression patterns of LEAFY/FLORICAULA and NEEDLY orthologs in female and male cones of the conifer genera Picea, Podocarpus, and Taxus: implications for current evo-devo hypotheses for gymnosperms. *Evolution and Development* **9**, 446–459.

Verbeke, J.A. (1992) Fusion events during floral morphogenesis. *Annual Review of Plant Physiology and Plant Molecular Biology* **43**, 583–598.

Villanneva, J.M., Broadhvest, J., Hauser, B.A., Meister, R.J., Schneitz, K. and Gasser, C.S. (1999) INNER NOOUTER regulates abaxial–adaxial patterning in Arabidopsis ovules. *Genes and Development* **13**, 3160–3169.

von Goethe, J.W. (1790) *Versuch die Metamorphose der Pflanzen zu erklären*. C.W. Ettinger, Gotha, Germany.

Walker, D.B. (1978) Post-genital carpel fusion in catharanthus-roseus (Apocynaceae). 4. Significance of fusion. *American Journal of Botany* **65**, 119–121.

Williams, E.G., Sage, T.L. and Thien, L.B. (1993) Functional syncarpy by intercarpellary growth of pollen tubes in a primitive apocarpous angiosperm, illicium-floridanum (Illiciaceae). *American Journal of Botany* **80**, 137–142.

Williams, J.H. and Friedman, W.E. (2002) Identification of diploid endosperm in an early angiosperm lineage. *Nature* **415**, 522–526.

Williams, J.H. and Friedman, W.E. (2004) The four-celled female gametophyte of Illicium (Illiciaceae; Austrobaileyales): implications for understanding the origin and early evolution of monocots, eumagnoliids, and eudicots. *American Journal of Botany* **91**, 332–351.

Winter, K.U., Becker, A., Munster, T., Kim, J.T., Saedler, H. and 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 of the United States of America* **96**, 7342–7347.

Winter, K.U., Saedler, H. and Theissen, G. (2002) On the origin of class B floral homeotic genes: functional substitution and dominant inhibition in Arabidopsis by expression of an orthologue from the gymnosperm Gnetum. *Plant Journal* **31**, 457–475.

Yalovsky, S., Rodriguez-Concepcion, M., Bracha, K., Toledo-Ortiz, G. and Gruissem, W. (2000) Prenylation of the floral transcription factor APETALA1 modulates its function. *Plant Cell* **12**, 1257–1266.

Yamada, T., Ito, M. and Kato, M. (2003) Expression pattern of INNER NOOUTER homologue in Nymphaea (water lily family, Nymphaeaceae). *Development Genes and Evolution* **213**, 510–513.

Yamaguchi, T., Lee, D.Y., Miyao, A., Hirochika, H., An, G.H. and Hirano, H.Y. (2006) Functional diversification of the two C-class MADS box genes OSMADS3 and OSMADS58 in Oryza sativa. *Plant Cell* **18**, 15–28.

Yamaguchi, T., Nagasawa, N., Kawasaki, S., Matsuoka, M., Nagato, Y. and Hirano, H.Y. (2004) The YABBY gene DROOPING LEAF regulates carpel specification and midrib development in Oryza sativa. *Plant Cell* **16**, 500–509.

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

Zanis, M.J., Soltis, D.E., Soltis, P.S., Mathews, S. and Donoghue, M.J. (2002) The root of the angiosperms revisited. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 6848–6853.

Zhang, P.Y., Tan, H.T.W., Pwee, K.H. and Kumar, P.P. (2004) Conservation of class C function of floral organ development during 300 million years of evolution from gymnosperms to angiosperms. *Plant Journal* **37**, 566–577.