Argument in the evo-devo debate: say it with flowers!

Serena Della Pina, Erik Souer and Ronald Koes*  
Department of Molecular Cell Biology, Graduate School of Experimental Plant Sciences, VU University, de Boelelaan 1085, 1081 HV Amsterdam, The Netherlands  
* To whom correspondence should be addressed. E-mail: ronald.koes@vu.nl

Abstract

A key question in evolutionary developmental biology is how DNA sequence changes have directed the evolution of morphological diversity. The widely accepted view was that morphological changes resulted from differences in number and/or type of transcription factors, or even from small changes in the amino acid sequence of similar proteins. Research over the last two decades indicated that most of the developmental and genetic mechanisms that produce new structures involve proteins that are deeply conserved. These proteins are encoded by a type of genes known as ‘toolkit’ genes that control a plethora of processes essential for the correct development of the organism. Mutations in these toolkit genes produce deleterious pleiotropic effects. In contrast, alterations in regulatory regions affect their expression only at specific sites in the organism, facilitating morphological change at the tissue and organ levels. However, some examples from the animal and plant fields indicate that coding mutations also contributed to phenotypic evolution. Therefore, the main question at this point is to what extent these mechanisms have contributed to the evolution of morphological diversity. Today, an increasing amount of data, especially from the plant field, implies that changes in cis-regulatory sequences in fact played a major role in evolution.

Key words: cis-Regulatory elements, evo-devo, flower evolution, gene expression, inflorescence, morphological novelty.

Introduction

Plant and animal species both display widely divergent morphologies regarding, for example, the arrangement of different organs on their bodies or the shape and size of various body parts, which are thought to result from evolution via mutation and selection. The Cambrian explosion (550 million years ago, mya) marks the appearance of most animal phyla that are known today (Wray et al., 1996), while land plants began their first diversification 450 mya (Kenrick and Crane, 1997). Flowering plants (Angiosperms) appeared only around 90–130 mya (Crane et al., 1995). Compared to animals, the evolution of flowering plants on land was relatively fast. Plants ‘invented’ and diversified new structures, such as flowers, over a ‘small’ lapse of time, making these relatively young events more tractable for genetic dissection. Since an ever-growing number of plant species is amenable to genetic analyses and transgenesis, plants offer excellent opportunities for research in evolutionary developmental biology (evo-devo).

This review summarizes current knowledge of the molecular basis of morphological changes during the course of evolution. We present general ideas and give examples of experimental data that support these theories. While some examples are taken from the animal field, the emphasis is on plant development as this represents our main research subject.

Lessons from the animal field

Although it is widely accepted that morphological variation between organisms arose from genetic alterations, the molecular details remain poorly understood. Initially it was assumed that species-specific characters resulted from species-specific
proteins. However, as the number of sequenced genes, proteins, and eventually entire genomes grew, it became clear that the genomes of organisms with very different morphologies are composed mostly of conserved genes (Miyamoto et al., 1987; Martin et al., 2010) and that genes controlling development are no exception.

Among the first examples were the homeotic HOX genes that specify the identity of (para)segments along the anterior–posterior axis of embryos ranging from insects to mammals (Mallo et al., 2010). They encode conserved homeodomain transcription factors and are arranged in clusters with a conserved gene order that correlates with their expression patterns along the anterior–posterior axis in early embryos (colinearity). Despite 700 million years of evolutionary separation, mammalian HOX proteins can still functionally replace their Drosophila homologues. For example, ectopic expression of the mouse HOXa5 protein or the Drosophila homologue SEX COMBS REDUCED caused similar defects in transgenic flies, including the transformation of antennae into T1 legs (Zhao et al., 1993), although HOXa5 plays a different role in mice (Aubin et al., 1999). Similarly, expression of chicken HOXb1 could fully rescue the defects of Drosophila labial mutants (Lutz et al., 1996). Comparable findings were made with non-HOX genes. For instance, EYELESS of Drosophila and its mouse homologue PAX6/SMALL EYE are both required for the development of eyes. Even though mouse and insect eyes have no obvious morphological similarity, expression of mouse PAX6 can drive the development of ectopic eyes in transgenic flies (Halder et al., 1995). A similar conservation was observed for extracellular proteins such as HEDGEHOG and WNT and their signalling pathways, which pattern (regions of) embryos and various organs (De Robertis, 2008).

These findings lead to the conclusion that the development of morphologically disparate animals is governed by a ‘toolkit’ of deeply conserved genes. This triggered the hypothesis, which had already been put forward early on (King and Wilson, 1975), that morphological divergence results primarily from alterations in gene expression patterns (Doebly and Lukens, 1998; Carroll, 2000, 2008). This idea was supported by comparative analyses showing, for example, that changes in the body plan of vertebrate and invertebrate species correlated with shifts in HOX gene expression patterns (Burke et al., 1995; Angelini and Kaufman, 2005; Heffer and Pick, 2013).

Gene expression patterns are governed by complex gene regulatory networks consisting of trans-acting (transcription) factors that bind to specific cis-acting DNA sequences in downstream genes to promote or inhibit their transcription. Binding sites usually cluster in small regions, known as enhancers, which promote transcription in specific cells. Genes often contain multiple enhancers—for expression on different sites—that can lie many kilobases away from the coding sequence. Genes that control body architecture, such as HOX genes, are highly connected nodes. They are expressed throughout development in complex patterns via an array of distinct enhancers and transcription factors and in turn they regulate a vast number of subordinate genes involved in a plethora of developmental processes. By contrast, transcription factors involved in terminal differentiation processes or ‘physiology’ (input–output genes) tend to form less connected nodes.

The modification of a gene expression pattern requires (some) rewiring of gene regulatory networks through changes in cis-regulatory elements (CREs) or upstream transcription factor proteins (Fig. 1). Alterations in highly connected transcription factors will simultaneously affect many other genes and developmental processes resulting in mostly deleterious pleiotropic effects. This may be one reason for the amazing conservation of insect and vertebrate toolkit proteins after more than 700 million years of separation. Changes in less-connected input–output transcription factors are more likely to be fixed because they cause less pleiotropic effects and are thought to account for the subtle phenotypic variation at the species level (Carrera et al., 2009; Bhardwaj et al., 2010). The complex expression patterns of toolkit genes governing development are often controlled by multiple enhancers that are active on different sites: alterations in an individual enhancer may alter the expression of the toolkit gene on one site, without affecting expression on other sites, and thus have less pleiotropic effects than alterations in the encoded protein. Based on such theoretical considerations, Carroll argued that the evolution of ‘anatomy’ and ‘physiology’ is essentially different and that alterations in CREs, rather than protein-coding sequences must be the major contributor to the evolution of form (Carroll, 2005, 2008).

Although the idea of morphological diversification via evolution of CREs is attractive, others criticized it for being premature and based on insufficient evidence (Hoekstra and Coyne, 2007). Moreover, there is also evidence for morphological changes via evolution of transcription factors. In arthropods, for example, the appearance of insects with only six (thoracic) legs but no legs on the abdomen is associated with a change in the HOX protein ULTRABITHORAX (UBX). This difference enables UBX to repress the expression of DISTALLESS in the abdomen, which is necessary for the development of appendages (Galant and Carroll, 2002; Ronshaugen et al., 2002). Similarly, the mammalian HOX11a protein acquired the ability to activate certain target genes, possibly in relation to the evolution of the extraembryonic membrane required for embryogenesis in utero (Lynch et al., 2008). Thus, studies with animals support the idea that body architecture evolved both via changes in CREs and changes in the protein sequence. However, a very limited number of studies exist that give solid support to either one or the other hypothesis. As illustrated in the following sections, plants offer excellent tools in order to study the molecular processes that drive the evolution of body architecture and shape.

Lessons from the plant field

During evolution, angiosperms developed specialized organs and tissues that have contributed to their diversity and success. According to fossil records, flower-like structures
Evolution of morphological diversity originated 160–147 mya (Frohlich, 2006). Already in 1970, Ohno and colleagues hypothesized a significant role for gene duplication in the evolution of biological complexity (Ohno, 1970). Land plants underwent several whole-genome duplications (Jiao et al., 2011) that facilitated their evolution by providing multiple copies of all pre-existing genes. Duplicated genes were released from the constraint of maintaining their original function and could therefore evolve in different ways, through mutations in either regulatory or coding regions. In the following discussion, we review data on the importance of CRE and protein changes in the evolution of flowering plants, taking examples from a variety of stages of the plant life cycle.

Before flowering: architecture of the vegetative plant body

Angiosperms show relatively little diversity during embryogenesis (Weijers and Jurgens, 2005) apart from the typical differences between monocots and dicots; most architectural differences arise later in development. After germination, the shoot apical meristem grows indeterminately and leaf primordia initiate at its periphery. Leaves can be simple, consisting of a single blade (Fig. 2A), or compound, consisting of several leaflets on the same petiole (Fig. 2B). Shoot apical meristem cells, which are located at the periphery of the meristem, downregulate KNOX transcription factors to...
enable differentiation (Vollbrecht et al., 1991; Long et al., 1996). In Arabidopsis, maize and Antirrhinum KNOX genes are repressed in the developing simple leaves by orthologous MYB proteins encoded by ASYMMETRIC LEAVES1, ROUGH SHEATH2, and PHANTASTICA (ARP genes), respectively. In species with lobed or compound leaves, such as Arabidopsis lyrata and Cardamine hirsuta, KNOX genes are reactivated during leaf development (Hay and Tsiantis, 2006; Piazza et al., 2010). Ectopic ZmKN1 expression in tobacco causes severely lobed leaves (Sinha et al., 1993), while in tomato it causes iterations of the compound pattern, resulting in the formation of super-compound leaves bearing thousands of leaflets (Hareven et al., 1996). In Arabidopsis, ectopic expression of distinct class 1 KNOX genes resulted in lobed leaves in all cases (Hay and Tsiantis, 2010), while silencing of the KNOX gene SHOOT MERISTEMLESS (STM) inhibited the lobing of leaves of Arabidopsis suecica (Piazza et al., 2010). Swapping KNOX genes or promoters between Arabidopsis thaliana and relatives with lobed leaves revealed that the divergent KNOX expression patterns result from differences in CREs (Piazza et al., 2010). In addition, promoter swaps between Arabidopsis and Cardamine hirsuta suggested that a change in the promoter of STM may be sufficient to determine the leaf complexity in the latter (Hay and Tsiantis, 2006). Taken together, these data show that class 1 KNOX proteins of different species are functionally similar and that changes in the CREs of KNOX genes were a key factor in the divergence of leaf shape.

A subclade of species within the Fabaceae forms compound leaves that do not express KNOX1 genes (Champagne et al., 2007). Leaf complexity is established in these species by distinct genes, such as UNIFOLIATA (UNI) and
**Evolution of morphological diversity**

Angiosperms display an amazing variation regarding the moment when they switch from vegetative growth to flowering (flowering time). Environmental and endogenous signals determine the moment that a plant undergoes this transition. Temperature and day length are the principal cues that trigger the onset of flowering in a particular season. Because these requirements are species dependent, some plants need a particular day length (photoperiod) in order to flower, while others are photoperiod insensitive. In addition, some species require exposure to low winter temperatures (vernalization), while summer annual plants do not. External and internal stimuli converge in the regulation of the main components of the long-distance florigen signal, FLOWERING LOCUS T (FT) and its parologue TWIN SISTER OF FT (TSF) (Andres and Coupland, 2012; Song et al., 2013). After their induction in leaves, FT and TSF proteins are loaded into the phloem and translocated to the apex where they interact with a bZIP transcription factor, FLOWERING LOCUS D (FD) via a 14-3-3 protein (Abe et al., 2005; Wigge et al., 2005; Taoka et al., 2011). These FT–FD and TSF–FD transcriptional complexes activate the expression of several floral pathway integrators: SUPPRESSOR OF CONSTANS1 (SOC1), SQUAMOSA PROMOTER BINDING PROTEIN LIKE (SPL), and Flower Meristem Identity (FMI) genes to promote flowering (Pose et al., 2012). Once FMI genes are transcribed, the plant is irreversibly committed to floral initiation, and FT and TSF become unnecessary (Corbesier et al., 2007).

In the long-day plant *Arabidopsis*, FT is activated by the zinc-finger transcription factor CONSTANS (CO), CO activity is regulated transcriptionally by light and the circadian clock, and post-translationally by light, in such a way that during short days CO protein is degraded, whereas in long days CO is stabilized and FT transcription activated (Turck et al., 2008). CO has an ancient origin as homologues that can functionally replace CO in *Arabidopsis* can be found even in the alga *Chlamydomonas reinhardtii*. Interestingly, *CrCO* is also regulated by day length and the clock in *Chlamydomonas* and is involved in rhythmic output processes (Serrano et al., 2009).

Clearly, FT-like proteins represent a universal flowering signal conserved during plant evolution. The diversification of seasonal control of flowering appears to be due to rewiring of the upstream network that controls FT expression and/or the downstream network for the output of FT signal. In rice, a short-day plant, the CO-homologue HEADING DATE 1 (HD1) activates the FT-like genes *HD3A* and *RICE FT-LIKE1*, but only in short-day conditions. The difference with *Arabidopsis* is that in long days HD1 is converted from an activator into a repressor of *HD3A* by a pathway that is
controlled by phytochrome B, which confers a robust regulation by day length and light quality (Ishikawa et al., 2005, 2009). In addition, EARLY HEADING DATE 1 (EHDI), a B-type response regulator, induces HD3A in short days independently of HD1. (Doi et al., 2004). In another short-day plant, Fragaria vesca (strawberry), long days induce expression of FvFT and (subsequently) FvSOCI, as in Arabidopsis. However, in strawberry plants FvSOCI activates a repressor of flowering that is homologous to TERMINAL FLOWER1 from Arabidopsis (Mouhu et al., 2013).

Sugar beet, a biennial that requires vernalization and long days for flowering, contains two FT paralogues. One of these (BvFT2) has a similar function as FT in Arabidopsis, whereas the other (BvFT1) is a repressor of flowering that is thought to be involved in the vernalization response (Pin et al., 2010). The divergence of BvFT1 and BvFT2 function involved changes in their CREs, because they are differentially expressed in beet, and changes in the encoded proteins, because BvFT2 induces and BvFT1 represses flowering when expressed in Arabidopsis (Pin et al., 2010).

In other species, the COFT system acquired new roles in entirely different developmental processes. In Populus trees, which are perennials, FT promotes the onset of flowering, as in Arabidopsis, and in addition suppresses the growth cessation and bud set in the autumn when days are shortening (Bohlenius et al., 2006). Interestingly, the critical day length for downregulation of FT, growth cessation, and bud set is shorter for varieties growing at moderate latitudes compared to those from northern latitudes. This adaptive change, which enables the timely preparation for winter, is associated with an altered CO regulation, because of which CO mRNA peaks several hours earlier after dawn in trees from southern populations compared with northern populations (Bohlenius et al., 2006). However, the genetic changes that caused the divergent CO expression remain to be established.

In potato, short days activate two FT paralogues: one (StSP3A) involved in the transition to flowering, and the other (StSP6A) involved in the transition to tuberization (Navarro et al., 2011). Since StSP6A can functionally replace FT in Arabidopsis and rice HD3A promotes both flowering and tuberization when expressed in potato, the functional divergence of StSP3A and StSP6A seems mostly due to changes in their transcriptional regulation and CREs rather than the encoded proteins (Navarro et al., 2011).

**Say it with flowers: the evolution of inflorescences**

After transition to flowering, flowers emerge in specific positions on the plant body. Some species form a single (solitary) flower, while others generate inflorescence branches that bear many flowers in a variety of arrays (Fig. 2C). In cymes flowers develop from apical meristems and inflorescence growth continues from a secondary lateral (sympodial) meristem formed at its flank, whereas in racemes the apical meristem maintains its indeterminacy and flowers develop from lateral meristems (Krizek and Fletcher, 2005).

In Arabidopsis, the expression of AP1 and LFY is largely restricted to floral meristems, although LFY is also expressed at a low level during vegetative growths. These genes are strongly expressed in lateral (floral) meristems, while in the apical (inflorescence) meristem their expression is prevented by TERMINAL FLOWER 1 (TFL1), in Arabidopsis, and CENTRORADIALIS in Antirrhinum (Bradley et al., 1996; Bradley et al., 1997). Constitutive expression of either LFY or AP1, which positively regulates each other in a regulatory loop, causes precocious flowering and converts the apical meristem into a flower, and, consequently, the racemose inflorescence into a solitary flower (Mandel and Yanofsky, 1995; Weigel and Nilsson, 1995). Hence, the spatiotemporal regulation of LFY and/or AP1 transcription is the major factor determining when and where flowers form in Arabidopsis. Similar results were obtained with other species, including trees such as aspen and citrus (Weigel and Nilsson, 1995; Pena et al., 2001). However in Petunia, a cyme, constitutive expression of LFY or the Petunia homologue ABERRANT LEAF AND FLOWER (ALF) does not affect flowering time or inflorescence architecture. This indicates that the transcription of ALF, although required for floral identity, is not the limiting factor that restricts flower formation in time or space. Instead, the limiting factor is another FMI gene: DOUBLE TOP (DOT) (Souver et al., 2008). DOT is the homologue of UNUSUAL FLORAL ORGANS (UFO) from Arabidopsis. DOT and UFO are functionally interchangeable F-box protein components of an SCF-type ubiquitin ligase complex that binds to ALF/LFY to promote its transcription activation potential (Chae et al., 2008; Souer et al., 2008). While expression of DOT is restricted to (incipient) flowers, UFO is expressed in many meristems, including the shoot apical meristem in embryos and seedlings. Indeed, constitutive expression of DOT or UFO in Petunia causes precocious flowering and changes the cymose inflorescence into a solitary flower, while in Arabidopsis no such effects are seen. Transgenic experiments in which promoter-reporter genes were swapped between Petunia and Arabidopsis indicated that the diversification of DOT and UFO expression was caused at least in part through differences in their CREs (Kusters, 2011).

Interestingly, LFY expression patterns vary even between species with similar inflorescence architectures. Brassicaceae such as Ionopsidium acaule (violet cress, Iac) and Laevensworthia crassa (Lcr) bear flowers in the axils of rosette leaves (‘rosette flowering’) and can be seen as racemes with reduced internodes (Shu et al., 2000). When introduced in Arabidopsis a LcrLFY transgene, or a fusion of the LcrLFY promoter (LcrLFYp) to the A LFY coding sequence, rectifies the lfy phenotype and in addition converts the apical meristem into a terminal flower, which fits with the finding that LcrLFYp is active in the apical inflorescence meristem (Yoon and Baum, 2004; Sliwinski et al., 2006). Even though IacLFY is expressed in floral meristems and the apical meristem in I. acaule, IacLFYp is not active in the apical meristem of Arabidopsis, and a IacLFY transgene does not affect the development the apical meristem in Arabidopsis (Shu et al., 2000; Yoon and Baum, 2004). These findings imply that LFY expression patterns vary even between closely related species through alterations in CREs.
within their promoters or in the upstream network without any obvious alterations in architecture. It may be that the apical inflorescence meristem of *L. crassa* and *I. acaule* do not acquire floral identity, even when it expresses LFY, because it lacks/has lost expression of an essential partner, such as the UFO/DOT homologue. The picture that emerges is that expression patterns of distinct FMI genes diverged substantially between angiosperms and thus created new patterns where essential partners are coexpressed and floral identity is established.

**Flowers of all hues: variations in flower architecture**

In angiosperms, LFY activates several genes involved in floral morphogenesis, including homeotic floral organ identity genes. Many of these encode MADS-box transcription factors which specify the fate of emerging organ primordia. According to the ABC model, which is based on phenotypes of homeotic *Arabidopsis thaliana* and *Antirrhinum majus* mutants, three classes of genes (A, B, and C) specify the identity of floral organ primordia (Coen and Meyerowitz, 1991). The A function specifies sepal identity in the outermost whorl (whorl 1), coexpression of A and B function specifies petal identity in whorl 2, B and C genes together control stamen identity in whorl 3, while the C function alone specifies carpel identity in whorl 4. In addition, the A and C functions antagonize each other. E function genes are a later addition to the model and act as coregulators in controlling the four floral organ identities (Theissen, 2001). Nowadays, a significant amount of data is available relating to the conservation/evolution of these transcription factors.

Flowers present an enormous variation in organ number, size, and shape (Fig. 2D–K). In gymnosperms, male and female sex organs develop on separate structures (cones) that lack most of the features seen in flowers. Putative B- and C-function orthologues can be found in gymnosperms as well: B-class genes are male specific, while C genes are expressed in both male and female organs (Tandre et al., 1998). These genes still show a significant level of functional conservation between gymnosperm and angiosperm despite millions of years of separation (Rutledge et al., 1998). This finding has been used as starting point to explain flower origin. According to the out-of-male (or out-of-female) theory (Theissen and Becker, 2004), class-B genes worked as a ‘switch’ to control female (or male) organ development in male (or female) cones. The formation of an apical–basal dissimilarity in B gene expression within the cone could have led to the hermaphroditic precursors of flowers. It is conceivable that both *trans*- and *cis*-regulatory element alterations are responsible for alterations in the expression pattern of these B genes. Further experiments are necessary to elucidate their evolution; however, these studies are hampered by the fact that basal angiosperms are not amenable to genetic dissection and that the ancestral plants bearing hermaphroditic precursors of flowers are unfortunately extinct.

*APETALA3 (AP3)* together with *PISTILLATA* belong to the B-type MADS box genes that are involved in the regulation of petal identity specification in *Arabidopsis* (Jenik and Irish, 2001). In Ranunculaceae, the AP3 lineage has undergone many duplication events, giving rise to three paralogous AP3 lineages (AP3-1, AP3-2, and AP3-3), which are found throughout the family (Kramer et al., 2003). Species bearing petals express AP3-3 orthologues specifically in petal primordia and developing petals, while species lacking petals lack AP3-3 expression (Zhang et al., 2013). AP3-3 expression seems to have been lost several times independently by either deletions or insertions in the coding, promoter, or intronic regions. The independent occurrence of apetalous species in the Ranunculaceae reflects a possible advantage: formation of petals and nectaries requires energy, which explains why many taxa have evolved general pollination systems with sepals, stamens, or filaments that are attractive to insects.

Another gene duplication, followed by a frameshift mutation in the C-terminal region of the AP3 gene, formed two new lineages in the core eudicots: euAP3 and TM6 (Vandenbussche et al., 2003). While *Arabidopsis* and *Antirrhinum* have lost their TM6 copy and maintained only euAP3 (*AtAP3* and *AmDEFICENS*), *Petrula hybrida* maintained both AP3 (PhDEF) and TM6. Mutation in PhDEF shows a conversion of petals to sepals, but the stamens are still made due to the fact that TM6 works redundantly in their specification: indeed, the double mutant transforms petals to sepals and stamens to carpels (Rijpkema et al., 2006). TM6 shows a third and fourth whorl expression pattern, while PhDEF is expressed in the second and third whorl. Interestingly, the divergence of the DEF/TM6 expression patterns correlates with the loss of a conserved promoter element in the TM6 lineage that might be responsible for the subfunctionalization of these genes. Ectopic expression of TM6 is capable of inducing petal development in a def mutant background, which shows that protein function is identical and that subfunctionalization results from changes in TM6 regulation (Vandenbussche et al., 2004; Rijpkema et al., 2006).

Heterotopic expression of B-function genes in whorl one of *Tulipa* explains the presence of petals in the outer whorl, also known as tepals (Kanno et al., 2003). Several species of the Solanaceae family show morphological alterations of whorl 1 too: inflated calyx syndrome is an amazing floral morphological novelty where, after pollination, the growth of sepals restarts thus giving rise to a balloon-like structure that encapsulates the mature berry. Research in *Physalis* has suggested that a change in the promoter region of MPF2—a member of the StMADS16 clade—accounts for the alteration in its expression, which is responsible for the inflated calyx syndrome in this species (He and Saedler, 2005).

An interesting exception to the conserved floral ground plan of eudicots is found in *Lacandonia schismatica*, where stamens occur in the centre of the flower and are surrounded by carpels. This different order in flower organs appears to be due to different spatiotemporal expression patterns in B- and C-function genes (Alvarez-Buylla et al., 2010), but further experiments are necessary in order to elucidate whether the new expression patterns arose via mutations in CREs...
of B and C genes or in the upstream regulatory network, or in both.

Flowers can display radial symmetry (Fig. 2I and J) if all petals develop and fuse equally or bilateral symmetry (zygomerphy) (Fig. 2H) if their petals develop in an unequal way. This difference in flower symmetry evolved multiple times during angiosperm evolution and it involves the expression of TEOSINTE BRANCHED1/CYCLOIDEA/PCF (TCP) genes along the dorsoventral axis. In radial asymmetric or zygomorphic flowers these genes are strongly expressed in the dorsal petals, while in symmetric or actinomorphic flowers they are present at too low level in order to effect flower symmetry. Yang et al. (2012) found that the promoters of two genes of the CYCLOIDEA2 (CYC) clade, CYC1C and CYCID of Primulina heterotricha, have consensus CYC-binding sites, resulting in positive autoregulation and cross-regulation of each other’s expression. This feedback loop seems to operate in many species which bear zygomorphic flowers, since consensus CYC-binding sites were identified in the promoters of their CYC2 genes. In contrast, the promoters of CYC2 genes in the actinomorphic lineage lack these CYC-binding sites and, consequently, the autoregulatory loop necessary to maintain their expression during flower development. It thus seems that a similar change in the CREs of CYC2 genes arose multiple times in angiosperms, which could explain the independent origin of floral zygomorphy.

**Terminal differentiation of epidermal cells**

Changes in the expression pattern also drove the evolution of metabolic pathways, such as plant pigmentation. Anthocyanins are responsible for most of the orange, red, and purple colouration in angiosperms (Fig. 2D–K) and serve a variety of (pleiotropic) functions: in vegetative tissues they protect from high light intensities, while in flowers and fruits they attract animals that aid pollination and seed dispersal (Buer et al., 2010).

Structural anthocyanin genes encoding the enzymes of the pathway are activated by a conserved complex consisting of a MYB, a basic helix–loop–helix (HLH), and a WD40 protein (MBW; Koes et al., 2005; Grotewold, 2006). Although the MBW complex is conserved between monocot and dicot species, its function has diverged substantially, in terms of downstream genes and processes that are activated. The MBW activates all structural genes from the anthocyanin pathway in maize, but only a subset of structural genes in most dicots. Exchanging either MYB and HLH regulators or gene promoters between species showed that this is due to differences in the structural gene promoters and not in the MYB and HLH proteins (Koes et al., 1994; Quattrocchio et al., 1998).

Phenotypes of Arabidopsis and Petunia mutants revealed that the divergence of MBW function extends to several other processes involved in the differentiation of epidermal cells (Ramsay and Glover, 2005). In Petunia, these include the division and morphogenesis of seed coat cells and several vacuolar processes that affect petal colouration, such as vacuolar acidification and stabilization of vacuolar anthocyanins. In Arabidopsis, MBW complexes specify the identity of incipient trichomes in stems and leaves and no-hair cells in the root epidermis and the production of mucilage by seeds. Interestingly, in Petunia or maize, no trichome defects are seen in gain- or loss-of-function mutants for MBW components, indicating that in these species the MBW complex has no role in trichome development. Nevertheless, the maize WD40 and HLH proteins PALE ALEURONE COLOR (PAC) and RED (R) can drive trichome development when expressed in Arabidopsis (Lloyd et al., 1992; Carey et al., 2004). This indicates that Arabidopsis MBW genes acquired their role in trichome development (and possibly no-hair cells in the root epidermis) via changes in the CREs of downstream trichome genes that brought them under MBW control. Genome-wide analysis recently identified some 20–40 target genes that are bound and regulated by GLABRA1 (GL1) and/or GL3, the MYB and HLH components of the Arabidopsis MBW complex (Morohashi and Grotewold, 2009). Comparative analysis with other species such as Petunia or maize may shed light on how these old genes learned new tricks.

Because pigmentation is an easy read-out of gene expression, it is a convenient system to use to study (adaptive) changes in gene expression patterns. The pigmentation of floral tissues together with other features, such as scent and flower morphology, attracts visits of specific animals (bees, moths, birds, bats) for pollination. Changes in any of these floral features (pollination syndrome) might lead to attraction of different pollinators and may, in theory, result in genetic isolation and ultimately speciation (Wessinger and Rausher, 2012). Petunia integrifolia, for example, has coloured flowers with a short wide tube that are visited by bumblebees, whereas Petunia axillaris has white scented flowers with a long narrow tube that are visited by nocturnal hawkmoths. Inactivation in Petunia axillaris of the ANTHOCYANIN2 gene, a MYB regulator, causes different petal pigmentation (Quattrocchio et al., 1999; Hoballah et al., 2007). Interestingly, an2 mutations were fixed in Petunia axillaris independently at least five times, but mutations in other anthocyanin genes were not. Presumably this is because AN2 expression is confined to the petal limb while other anthocyanin genes are also expressed on many other sites, which is the reason why mutations in AN2 are the least pleiotropic (Quattrocchio et al., 1999). A shift from bee to hummingbird pollination is associated with transition from blue-purple to red coloured flowers, together with other morphological alterations. The most frequent cause of these colour changes is a reduced expression of Flavonoid 3′-Hydroxylase (F3′H) and/or F3′5′H (Wessinger and Rausher, 2012). Expression of F3′5′H, which is primarily required for flower pigmentation, was abolished by mutations in either the coding sequence or a CRE. By contrast in F3′H, which is involved in other processes besides flower pigmentation, mutations occurred predominantly in its CREs, thus reducing F3′H expression mainly in flowers (Wessinger and Rausher, 2012).

The cases summarized above concern loss-of-function mutations that reduce or eliminate pigmentation in certain tissues. They all have an impact on elements that cause relatively few pleiotropic effects, either a CRE or the coding sequence.
of a parologue with a restricted expression pattern. Cases in which novel pigmentation patterns were acquired are more scarcely documented. That may be because such mutations are not favoured and/or they are more difficult to prove without knowing the ancestral state. A major part of the colour difference between the pale red bumblebee-pollinated flowers of *Mimulus lewisii* and the bright red hummingbird-pollinated flowers of *Mimulus cardinalis* is due to the difference in expression level of a MYB protein, ROSE INTENSITY 1, which inhibits the activity of the MBW complex (Yuan et al., 2013). The different expression levels appear to be due to an alteration in ROI CREs, but since the ancestral state is unknown it is unclear whether this involved the loss of an activating CRE in *Mimulus cardinalis* or a gain in *Mimulus lewisii*.

The creation of new pigmentation patterns by gain-of-function mutations has been documented in a few domesticated crops. In all species analysed, the WD40 partner of the MBW complex is more or less ubiquitously expressed and the pigmentation pattern is largely dictated by the expression patterns of the MYB and basic HLH partners (Koes et al., 2005). Ectopic expression of the MYB protein is sufficient to establish the pigmentation of new tissues (Spelt et al., 2000; Nesi et al., 2001). The appearance of varieties with new pigmentation patterns was indeed associated with alterations in the regulatory region of the MYB anthocyanin regulators. For instance, the appearance of blood oranges with anthocyanins in the flesh of the fruit happened independently on at least two occasions; in both cases, new CREs originated from a retrotransposon insertion into the RUBY gene, which encodes a MYB component of the MBW complex (Butelli et al., 2012). The colour difference between red and green apples can also be traced to a CRE in the *MYB10* allele of red apples that results in enhanced *MYB10* expression (Espley et al., 2007). This novel CRE originates from a microsatellite and creates a new binding site for MYB10, itself resulting in an autoregulatory loop (Espley et al., 2009).

**Conclusion**

In this review, we have tried to elucidate some of the general mechanisms that have been at work throughout plant and animal evolution. The appearance of novel structures and their subsequent diversification relied primarily on alterations in spatiotemporal gene expression patterns, rather than the sudden appearance of completely new genes. Indeed, in most cases morphological novelty originates from changes in the timing, rates, or pattern of expression of pre-existing genes. To our knowledge, very few cases have been documented where a new gene expression pattern could be attributed to an alteration in a transcription factor protein. That might be because such events are rarely fixed during evolution, due to their pleiotropic defects, and/or because such events are more difficult to identify and prove experimentally. However, in a substantial number of cases, the cause of an altered gene expression pattern has been shown to result from changes in gene promoters, although the specific CREs and DNA changes involved have been traced down in only a few cases. This may be due to the fact that CREs can lie many kilobases away from the coding sequence and are difficult to recognize since single nucleotide changes may be sufficient to create or remove transcription factor-binding sites. The general picture is that most changes, including the few in transcription factor proteins, seem to have been selected for minimum pleiotropic effects.

One could envisage several other mechanisms -in addition to the modification of tissue-specific CREs involved in transcription control -to rewire gene regulatory networks with minimal pleiotropic effects. For example, alternative splicing (Syed et al., 2012) or polyadenylation may lead to the expression of novel forms of a protein, with new functions while leaving the old function unharmed. Alterations in (trans-acting) miRNAs, (cis-acting) complementary sequences within mRNAs (Kosik, 2009), DNA methylation (Cortijo et al., 2014), and histone modifications (Turck and Coupland, 2013) could, theoretically, modify gene expression patterns in similar subtle ways as changes in CREs. To what extent such mechanisms contribute to morphological diversification remains to be determined.

**Acknowledgements**

This work was supported by a grant from the Netherlands Organization for Scientific Research (NWO). We apologize for not including many of the exciting studies and references due to space limitations.

**References**


