Molecular Analysis of the anthocyanin2 Gene of Petunia and Its Role in the Evolution of Flower Color

Francesca Quattrocchio, John Wing, Karel van der Woude, Erik Souer, Nick de Vetten, Joseph Mol, and Ronald Koes

Department of Genetics, Institute for Molecular Biological Sciences, Vrije Universiteit, BioCentrum Amsterdam, de Boelelaan 1087, 1081 HV Amsterdam, The Netherlands

The shape and color of flowers are important for plant reproduction because they attract pollinators such as insects and birds. Therefore, it is thought that alterations in these traits may result in the attraction of different pollinators, genetic isolation, and ultimately, (sympatric) speciation. Petunia integrifolia and P. axillaris bear flowers with different shapes and colors that appear to be visited by different insects. The anthocyanin2 (an2) locus, a regulator of the anthocyanin biosynthetic pathway, is the main determinant of color differences. Here, we report an analysis of molecular events at the an2 locus that occur during Petunia spp evolution. We isolated an2 by transposon tagging and found that it encodes a MYB domain protein, indicating that it is a transcription factor. Analysis of P. axillaris subspecies with white flowers showed that they contain an2 alleles with two alternative frameshifts at one site, apparently caused by the insertion and subsequent excision of a transposon. A third an2 allele has a nonsense mutation elsewhere, indicating that it arose independently. The distribution of polymorphisms in an2 alleles suggests that the loss of an2 function and the consequent changes in floral color were not the primary cause for genetic separation of P. integrifolia and P. axillaris. Rather, they were events that occurred late in the speciation process, possibly to reinforce genetic isolation and complete speciation.

INTRODUCTION

Flowers are the structures containing the male and female sex organs of angiosperms. Flowers of diverse species display a wide range of different morphologies and pollination strategies. For instance, flowers of wind-pollinated species usually possess small and inconspicuous petals or no petals at all, whereas flowers of insect-pollinated plants usually possess large, brightly colored, and patterned petals that serve as visual signals and a landing site for visiting insects.

Recent experiments suggest that the wide variety of plant and flower morphologies may have depended on the evolution of a relatively small number of genes. First, mutations at single loci, usually isolated by breeders or researchers, are sufficient to cause fundamental alterations in inflorescence architecture (Doebley et al., 1997; Souer et al., 1998). Similarly, the different shapes, colors, and color patterns of naturally occurring Mimulus (monkeyflower) spp are due to alterations at only a few (major) loci (Bradshaw et al., 1995).

Second, even very different inflorescence and flower architectures appear to be determined by genes that often encode conserved proteins but that differ in their expression patterns (reviewed in Doebley and Lukens, 1998).

The isolation of key regulatory loci and analysis of the molecular alterations that have taken place in them will provide new insights into the evolution and diversification of flower morphology. The biosynthesis of anthocyanin flower pigments is particularly suited for such studies, because it is a well-defined biochemical pathway that is being studied simultaneously in distinct species with different flower morphologies and pollination strategies, such as maize (wind pollinated with flowers that lack petals), Arabidopsis (self-pollinating with white flowers), Antirrhinum (insect pollinated with colored flowers), and Petunia spp (both white and colored flowering species that are visited by different insects) (reviewed in Martin and Gerats, 1993; Holton and Comish, 1995; Mol et al., 1998).

It is generally believed that the structural genes encoding the enzymes of the anthocyanin pathway are similar in all angiosperm species (Holton and Cornish, 1995). Recently, however, it was shown that one of the last steps in the pathway, glutathionation, is catalyzed by glutathione S-transferases of different evolutionary origin in maize and Petunia spp (Marrs et al., 1995; Alfenito et al., 1998). Moreover, the
expression of the structural genes appears to be regulated differently in distinct species (reviewed in Mol et al., 1998; Weisshaar and Jenkins, 1998). Mutational analyses showed that in maize, transcription of the entire set of structural anthocyanin biosynthesis genes is controlled as a single unit by two families of regulatory genes. The so-called c1/pl and the r/b gene families comprise multiple paralogous genes that encode functionally similar proteins that include a MYB domain and a basic helix-loop-helix (bHLH) domain, respectively, and that have distinct expression patterns (Ludwig and Wessler, 1990; Cone et al., 1993).

In Antirrhinum and Petunia spp flowers, however, the pathway appears to be regulated in at least two distinct units (Martin et al., 1991; Quattrocchio et al., 1993). Mutations at the anthocyanin1 (an1), an2, an4, and an11 loci of Petunia hybrida cause inactivation of structural genes acting late in the pathway in specific parts of the flower (Beld et al., 1989; Quattrocchio et al., 1993; Huits et al., 1994a). The same holds true for mutations at delila (del) and rosea of Antirrhinum (Martin et al., 1991). Remarkably, the early structural genes remain transcriptionally active in these Petunia spp and Antirrhinum mutants, suggesting that they are activated by another set of regulatory genes (Moyano et al., 1996; Quattrocchio et al., 1998).

These findings can be explained in at least two ways: (1) the regulators of flower pigmentation in Antirrhinum and Petunia spp are not related to c1 and r from maize; or (2) c1- and r-related genes do control flower pigmentation, and regulatory differences are due to divergent evolution of the target structural genes. Although several regulatory anthocyanin genes have now been isolated from Antirrhinum and Petunia spp, the data are too incomplete to exclude either of these two possibilities.

Molecular analysis of the an11 locus of Petunia spp showed that it encodes a WD-40 repeat protein that is highly conserved in nature, even in animals and yeast that cannot synthesize anthocyanins (de Vetten et al., 1997). However, no information is available on the function of these an11 homologs in their cognate hosts. The an1 locus includes a gene that encodes a bHLH protein but this gene cannot be considered to be an ortholog of the maize r gene (C. Spelt and R. Koes, unpublished data), because another petunia gene, jaf13, has greater similarity to r (Quattrocchio et al., 1998). Although jaf13 can replace the function of r in some functional assays, it cannot complement r- mutants (Quattrocchio et al., 1998). Similarly, the del gene of Antirrhinum, encoding a bHLH protein with homology to R (Goodrich et al., 1992), failed to complement r- mutants (Mooney et al., 1995). Therefore, it is not clear whether jaf13, del, and r are truly orthologous. However, our knowledge of the factors that control anthocyanin biosynthesis is far from complete because, for instance, information on the products encoded by an2 and an4 of Petunia spp and rosea of Antirrhinum is lacking. These loci are particularly interesting because they appear to have played a role in the diversification of flower color and patterns in nature.

P. integrifolia and P. axillaris subspecies occur naturally in South America in partially overlapping areas. Yet, both species remain genetically separated, apparently because their flowers are visited by different insects (Wijsman, 1982, 1983). P. integrifolia flowers have a purple corolla with a short wide tube and are pollinated by bees (Figure 1A), whereas P. axillaris flowers have a white corolla with a long narrow tube (Figure 1B), which is typical of moth-pollinated flowers (Wijsman, 1982). Manual cross-pollination, however, readily produces fertile progeny. In fact, the garden petunia (P. hybrida) is thought to be derived from such interspecific crosses (Wijsman, 1983; Sink, 1984; Koes et al., 1987).

Genetic analyses showed that differences at five loci, an2, an4, hydroxylation-at-five (hf1), flavonols (fl), and pollen (po), are responsible for the different floral organ colors of various subspecies of P. integrifolia (genotype An2+An4+hf1+Fl+Po+) and P. axillaris (genotype An2- an4- hf1 hf1-1Fl-1Po- or an2- an4- hf1-1Fl-1Po-) (Wijsman, 1983). The an4- and po- markers of P. axillaris are responsible for the yellow anthers, but because these markers do not affect coloration of the corolla, they are not discussed here (cf. van Tunen et al., 1991; Quattrocchio et al., 1993). The markers an2-, hf1-1, and Fl- contribute in different degrees to the white color of the P. axillaris flower corolla. The an2- genotype strongly reduces but does not completely abolish the transcription of structural anthocyanin genes and coloration, when compared with An2+. On their own, neither Fl+ nor the partially

Figure 1. Flower Phenotypes of Petunia spp.
(A) Flowers of P. integrifolia subsp violacea line S9.
(B) Flowers of P. axillaris subsp axillaris line S1.
(C) P. hybrida flower harboring the unstable an2-W82 allele.
(D) P. hybrida flower harboring a revertant allele derived from an2-W82.
active hfl-1 allele contributes strongly to the white flower color. Fh-1 makes the flower more bluish (copigmentation) without reducing the accumulation of delphinidin-type anthocyanins. Mutation of Hf1 results in the formation of redish cyanidin-type anthocyanins instead of delphinidin-type anthocyanins. Only the combination of Fh-1 and hfl-1 (or hfl-1) reduces the synthesis of anthocyanins; for unknown reasons, flavonol synthesis competes with cyanidin synthesis and not so much with delphinidin synthesis (Wiering and De Vlaming, 1984). During the evolution of petal color, the role of the an2 locus appears to have been more prominent than the combination of hfl and Fh-1. This is because (1) P. axillaris accessions can be either Fh-1 or hfl-1 (Wijsman, 1983), indicating that this difference arose late, after the genetic separation from P. integrifolia; (2) An2+* hfl-1 Fh-1 flowers can still accumulate a considerable amount of anthocyanins (see, e.g., Figures 6H and 6I in Quattrocchio et al., 1998), whereas (3) the effect of an an2- allele in a Hf1*Fh-1 background is considerably stronger (see Figures 6J and 6K in Quattrocchio et al., 1998).

To gain further insight into the molecular mechanisms that control flower pigmentation and into the evolution of those mechanisms, we conducted a molecular analysis of the an2 locus. We isolated an2 by transposon tagging and showed that it encodes a MYB domain protein. A comparison of the an2 alleles in Petunia spp showed that the lack of an2 function in P. axillaris flowers results from two independent loss-of-function mutations, one of which was apparently induced by the insertion and excision of a transposable element. Analysis of an2 polymorphisms indicated that mutations in an2 occurred relatively late in the speciation process, suggesting that in this instance, floral color change was not the primary speciation event but rather a mechanism to reinforce genetic separation.

RESULTS

Molecular Analysis of the an2 Locus

P. hybrida line W82 in the Amsterdam collection harbors an unstable an2 allele, herein named an2-W82, that is presumably identical to the an2-n allele described earlier (Cornu, 1977). The flowers of an2-W82/an2-W82 plants have pale-colored corolla limbs with sectors and spots of a different intensity (Figure 1C). Among progeny obtained by self-pollination of an2-W82 homozygotes, we identified four plants with full-colored flowers (Figure 1D) that originated from independent sporogenic reversion events. Because most unstable alleles in P. hybrida contain insertions of the 284-bp transposable element dTph1 (van Houwelingen et al., 1998), we anticipated the presence of dTph1 in the an2-W82 allele. To clone a fragment of the an2 locus, we isolated dTph1 flanking sequences that were present in two an2-W82 homozygous plants, but not in a homozygote for the revertant allele An2*+R1, by a combination of inverse polymerase chain reactions (PCRs) and differential screening of cloned amplification products (Souer et al., 1995). One of these dTph1 flanking sequences, jaf41, hybridized with a 2.2-kb fragment in an2-W82 homozygous plants, whereas a 1.9-kb fragment was detected in plants harboring four independently derived excision alleles (Figure 2A). This result showed that jaf41 contains part of the an2 locus. We subsequently isolated the complete locus and cDNA clones by hybridizing jaf41 with corresponding libraries that were made from the An2+ P. hybrida line V26.

The sequence of an an2 cDNA revealed a large open

![Figure 2](image-url)
reading frame (ORF) of 765 bp (Figure 2B) that is preceded by a small (69 bp) ORF. In an2-W82, the dTph1 insertion duplicated 8 bp, including the ATG start codon of the large ORF (Figure 2C), without lowering the an2 mRNA level (data not shown). This suggests that the mutant phenotype is caused by reduced translation initiation of the large ORF and implies that this ORF encodes AN2. The finding that naturally occurring an2 alleles contain mutations that disrupt the 765-bp ORF lends further support to this conclusion (see below).

Database searches revealed that the AN2 protein is similar to a range of so-called MYB domain proteins from animals and plants. Parsimony analysis showed that, in general, AN2 is more related to plant MYB proteins than to the animal c-MYB, the prototype of this superfamily (Figure 3A). The protein with the highest similarity found in the databases is that encoded by the myb75 gene of Arabidopsis, a gene whose function is not known (Kranz et al., 1998). Because the majority of genes in this Arabidopsis family have been sequenced, this suggests that myb75 may be the Arabidopsis homolog of an2. However, AN2 does not display significantly greater similarity to other MYB domain proteins implicated in flavonoid biosynthesis, such as MYB305 and MYB340 of Antirrhinum (Moyano et al., 1996), C1 and P from maize (Paz-Ares et al., 1987; Grotwold et al., 1991), or MYBPH3 from P. hybrida (Solano et al., 1995), than it does to MYB domain proteins involved in other processes (Figure 3B). In all cases, the similarity was restricted to the two repeats that make up the MYB domain, whereas no similarities were detectable in the C-terminal half of these proteins (Figure 3B). Comparison of the genomic and cDNA sequences showed that the an2 gene is split by two introns (Figure 2B) in positions that are conserved in other myb-related genes.

Polymorphisms among an2 Alleles

To investigate the evolutionary relationships between the various An2+ and an2− alleles of different Petunia spp, we conducted DNA gel blot analyses. These experiments showed that P. axillaris subsp and an2− P. hybrida lines contain genomic an2 fragments with nearly identical restriction maps, whereas the P. integrifolia subsp and An2+ P. hybrida lines harbored different polymorphic fragments (data not shown). We also detected polymorphisms by using PCR amplification of an an2 fragment that spans both introns (Figure 4A). In an2− lines of P. axillaris or P. hybrida, this amplification generated a 1.2-kb fragment, whereas the same fragment in An2+ lines of P. integrifolia and P. hybrida varied in size between 1.2 and 1.4 kb (Figure 4A). Further mapping with different primer combinations showed that the polymorphism resides in the second an2 intron (data not shown). To determine the nature of the polymorphism, we PCR amplified and sequenced the second intron of the an2 allele of P. axillaris S1 and P. integrifolia S9 and S12. Comparison of these sequences showed that the polymorphism resulted from two insertions or deletions—one of 36 bp and a second of variable length (≈50 bp in P. integrifolia S12 and 122 bp in P. integrifolia S9) (Figure 4A). A closer examination of these insertions/deletions and surrounding region did not reveal specific features (e.g., repeated sequences) that might account for a mechanism by which the polymorphism was established.

Figure 3. an2 Encodes a MYB Domain Protein.

(A) A neighbor-joining phylogenic tree of AN2 and selected MYB domain proteins from other species. The numbers in the branches indicate the percentage of bootstrap support after 500 replicates. For construction of the tree, we used only the MYB domains; for c-MYB, we also excluded repeat 1 of its MYB domain, because this repeat is not found in the plant proteins.

(B) Alignment of the MYB domain of AN2 and proteins controlling anthocyanin (C1) and phlobaphene (P) synthesis in maize, phenylpropanoid metabolism (MYB305) in Antirrhinum and Petunia spp (MYBPH3), cell shape in Antirrhinum (MIXTA) and Petunia spp (MYBPH1), trichome development in Arabidopsis (GL1), and hematopoiesis in humans (c-MYB) (reviewed in Martin and Paz-Ares, 1997), and MYB75 of Arabidopsis, a protein with unknown function. Residues on a black background denote sequence similarity between AN2 and one or more of the other proteins.
Alignments of the an2 mRNA sequences also showed that the an2− and An2+ alleles fall in two separate groups (Figures 4B and 5). The phylogenetic trees that were obtained in this way were fully consistent with the distribution of restriction fragment length polymorphisms and intron size polymorphisms described above.

Taken together, these data indicate that An2+ and an2− alleles comprise two distinct groups that have been maintained in genetically separated populations, during which time the observed polymorphisms (restriction fragment length polymorphisms, the intron size polymorphisms, and sequence polymorphisms in the coding region) arose. The high degree of similarity between an2− alleles indicates that the an2− alleles found in P. hybrida are introgressed an2− alleles of P. axillaris.

**Natural an2− Alleles Contain Frameshift and Nonsense Mutations**

To determine the nature of the mutation that inactivated the an2− alleles, we determined the an2 mRNA level in petals of different Petunia spp lines. Because an2 mRNA was of too low abundance to detect by RNA gel blot analysis, we used a quantitative reverse transcription (RT)-PCR assay. As an internal control, we coamplified products of the chsA gene, encoding the enzyme chalcone synthase that catalyzes the first step in flavonoid synthesis. Even though chsA is expressed independently of an2, its temporal expression pattern during petal development is very similar to that of the an2-controlled genes involved in anthocyanin synthesis (Quattrocchio et al., 1993, 1998). Figure 5 shows that all the an2− Petunia spp lines tested express an2 transcripts in amounts that are only slightly lower than An2+ lines do. Given that these lines represent different species, the relatively small variations in an2 transcript abundance most likely result from differences in genetic background and/or decreased mRNA stability due to a blocked translation (see Discussion).

To determine the structure of the an2 transcripts detected in An2+ and an2− lines, we cloned the RT-PCR products and determined their sequence. In the an2− alleles of P. axillaris X504 and P. hybrida W44, W22, and W59, we found a 4-bp insertion after codon 127, just downstream of the

**Figure 4.** Phylogenetic Relationships of an2 Alleles. (A) Polymorphic an2 fragments amplified from P. integrifolia (P.i.), P. axillaris (P.a.), and P. hybrida (P.h.) lines. The lane marked H20 is a control in which water without template DNA was used for amplification. The an2 genotype of the lines used is indicated by + (An2+), − (an2−), and m (an2 mutable). The positions of the two length markers are indicated at left. The diagram below shows the structure of the an2 gene as in Figure 1A, the positions of the PCR primers (arrowheads), and the insertions/deletions (open triangles under the gene) responsible for the polymorphisms in the amplification products. (B) A phylogenetic tree constructed by alignment of complete an2 protein coding sequences using the UPGMA (for unweighted pair method using arithmetic averages) algorithm. The thick bars indicate the standard error in the positions of the branch points. The sequences have the GenBank accession numbers AF146702 (P. hybrida V26), AF146703 (P. integrifolia S9), AF146704 (P. integrifolia S6), AF146705 (P. hybrida W115), AF146706 (P. hybrida W22), AF146707 (P. hybrida W44), AF146708 (P. axillaris S1), and AF146709 (P. axillaris S7).

**Figure 5.** Analysis of an2 Transcripts in an2− and An2+ Petunia spp Lines. Transcripts of the genes an2 and, as control, chsA were detected by quantitative RT-PCR.
conserved MYB domain. This insertion results in a reading frame shift and a premature stop codon, thereby preventing the translation of the nonconserved C-terminal half of the AN2 protein (Figures 6A and 6B). In the an2 alleles in P. axillaris S1 and X508 and P. hybrida W115, we found a 1-bp deletion in codon 127, which also prevents translation of the C-terminal half of AN2 (Figure 6B). Partial sequencing of the an2 allele of P. axillaris S8 showed that this allele contained the same 1-bp deletion.

Partial sequencing of the an2-S7 mRNA showed that the reading frame was intact at codon 127. We subsequently determined the complete mRNA sequence. This showed that an2-S7 contains a nonsense mutation farther downstream at codon 196 (Figure 6B).

The Frameshift and Nonsense Mutations Are Responsible for an2 Inactivation

We considered the possibility that the frameshift and nonsense mutations arose in an2 alleles that had been already inactivated by other mutations. It is very unlikely that mutations upstream of the frameshift had inactivated an2 prior to the frameshift, because the truncated protein encoded by an2-W22 has the same sequence as the corresponding region in that encoded by An2-V26 (Figure 6A). However, we could not exclude that inactivation had occurred downstream from the frameshift.

To identify evolutionary old mutations that might have inactivated the an2 alleles before the frameshift and nonsense

Figure 6. Analysis of the Mutations That Inactivate an2 Alleles.

(A) Alignment of translated An2 and an2 alleles. To produce the translations of the an2 alleles, the frameshift or nonsense mutations were ignored. The positions of these frameshift and nonsense mutations in an2 alleles are indicated by a dot on a red background. Polymorphisms specific for An2 and an2 alleles are indicated in purple and green, respectively. Other polymorphisms are highlighted in black. Numbering of the amino acid residues is indicated above the sequences (dots indicate every tenth residue). For species abbreviations, see the legend to Figure 4A.

(B) DNA and deduced protein sequences around the frameshift and nonsense mutations in an2 alleles. Codons 127 and 196, at which frameshift and nonsense mutations were found, are in red. The broken lines denote the sequences in between these two regions. Stop codons that terminate the AN2 protein coding sequence are boxed.

(C) Structure and activity of An2, an2, and hybrid alleles. The maps denote the mRNA structure of the different alleles. Black regions denote the MYB domain, and gray regions denote other translated sequences. White regions indicate mRNA sequences that are not translated due to the presence of a frameshift (an2-W44) or a nonsense mutation (an2-S7) at the site marked by a red asterisk. The green and purple bars indicate the position of polymorphisms, as given in (A). The activity of these alleles, when expressed from the 35S promoter, was measured after microprojectile delivery into an2-W115 petal cells by using a dfrA–luc reporter gene. Activities are given as the mean ±SE of five independent bombardment assays and are expressed in arbitrary units.
Flower Color Evolution in Petunia 1439

mutated, we searched for sequence polymorphisms that are consistently present in (nearly) all an2 alleles. Therefore, we aligned the translated AN2 sequences while ignoring the frameshift and nonsense mutations in the an2 alleles (Figure 6A). This revealed some sequence variations for single alleles that are scattered throughout the protein (Figure 6A). This distribution pattern indicates that these mutations arose relatively late in evolution, after the frameshift and nonsense mutations, and therefore were not responsible for gene inactivation in nature. Because the an2 alleles are no longer under selection pressure, one might expect that more such mutations will accumulate at a high rate.

At 13 positions, all clustered in the C terminus, the proteins of the An2- and an2 alleles differed consistently (Figure 6A), indicating that these variations occurred soon after the genetic separation of both groups of alleles. To examine whether these early alterations resulted in an inactive AN2 protein, we used a conserved ClaI restriction site to replace the last 65 codons of the wild-type An2-V26 cDNA with those from an2-W44 and fused the cDNAs to the constitutive 35S promoter of the cauliflower mosaic virus. To assay the activity of the various an2 constructs, we used a complementation assay in which they were delivered by particle bombardment into an2 petal cells. To measure AN2 activity, we assayed the expression of a codelivered reporter gene (lac, encoding luciferase) driven by the an2-responsive promoter of the dfrA gene, which encodes dihydroflavonol 4-reductase, a “late” enzyme in the anthocyanin pathway. To normalize the dfrA-driven luc activity for variations in transformation efficiency, we measured the activity of a codelivered reporter gene consisting of the β-glucuronidase coding sequence driven by the 35S promoter (35S-gus). Figure 6C shows that an2-W115, containing a −1-bp frameshift at codon 127, is a null allele because it fails to induce the dfrA-luc reporter gene. The hybrid allele An2-V26/W44, however, activated the dfrA promoter at least as efficiently as the wild-type An2-V26 allele. This indicates that the polymorphisms near the former stop codon of an2-W44 or its progenitor allele did not inactivate the protein. This apparent high tolerance for sequence variations cannot be taken as evidence that this part of the AN2 protein is without function, because truncation of the C-terminal domain, as in an2-S7, results in complete loss of AN2 activity (Figure 6C).

Taken together, these data strongly suggest that inactivation of an2 in nature was caused by the frameshift and nonsense mutations, and not by the sequence variations observed in the C terminus of the protein.

**DISCUSSION**

To elucidate the molecular mechanisms that control flower pigmentation patterns, we studied a set of regulatory genes that control the tissue-specific transcription of the structural genes encoding enzymes of the anthocyanin pathway. Early experiments indicated that transcription of these structural genes in *P. hybrida*, *Arabidopsis*, and tobacco could be activated by combined expression of the c1 and r genes from maize (Lloyd et al., 1992; Quattrocchio et al., 1998). However, this does not necessarily mean that r- and c1-homologous genes control pigmentation of the flower in vivo. In fact, the finding that an11 and an1 encode, respectively, a WD-40 protein (de Vetten et al., 1997) and a bHLH protein (C. Spelt and R. Koes, unpublished data) that had not been previously identified in other species indicates a different direction. Here, we report the isolation of an2, a third regulator of the anthocyanin pathway in Petunia spp, which acts in concert with an1 and an11 to activate transcription of structural anthocyanin genes in the petal limb.

**an2 Encodes a MYB Domain Protein**

The homology of AN2 to various MYB domain proteins suggests that AN2 is a transcription factor that activates a subset of structural anthocyanin genes. Such an activity would be consistent with the results of transient expression assays, which show that AN2 can activate promoter activity of the dfr gene (Figure 6C and Quattrocchio et al., 1998). Whether this activity involves binding of AN2 to the promoter of the structural genes or operates via an intermediate regulatory gene is currently under investigation.

Strikingly, AN2 does not display significantly more similarity to MYB305 and MYB340, two regulators of early structural genes in Antirrhinum (Sablowski et al., 1994; Moyano et al., 1996), or C1 and PI from maize. This finding, together with the observation that an2, myb305/myb340, and c1 control different sets of target structural genes in their respective hosts (reviewed in Mol et al., 1998), may at first sight suggest that these genes are not homologs. However, subsequent expression assays and complementation experiments showed that AN2 can replace C1 and vice versa, indicating that c1 and an2 are homologous genes (Quattrocchio et al., 1998). The different sets of target anthocyanin biosynthesis genes controlled by an2 and c1 in their cognate hosts therefore appear to be due to divergent evolution of the target genes rather than of the regulatory proteins (Koes et al., 1994; Quattrocchio et al., 1998).

For C1, it was shown that its specificity resides in the DNA binding MYB domain (Sainz et al., 1997; Williams and Grotewold, 1997) and that this domain interacts with bHLH proteins encoded by the r gene family (Goff et al., 1992). However, also in this more limited domain, the similarity between AN2 and C1 does not appear to be significantly higher than it is between functionally unrelated MYB domain proteins (Figure 3), even though AN2 is the only MYB protein known that can complement mutations in c1 and interact with maize R proteins or the petunia homolog J A F13 in transient expression assays (Quattrocchio et al., 1998) and in a yeast two-hybrid assay (A. Koorn and R. Koes, unpublished data). The C-terminal half of AN2 functions in yeast as a
transcriptional activation domain (A. Kroon and R. Koes, unpublished data), similar to the C-terminal domain of C1 (Goff et al., 1992). This suggests that the similarity between C1 and AN2 is higher than can be detected by sequence alignments or, in other words, that a significant number of alterations are tolerated—especially in the C-terminal domain—without loss of activity or specificity. This may account for the relatively high sequence divergence in the (former) C-terminal domains of an2− and An2− alleles (Figure 6A). Nevertheless, this domain appears essential for activity, because the truncated protein encoded by an2-S7 is inactive (Figure 6C).

In P. hybrida, mutations in either of the regulators an1 or an11 leads to a complete loss of expression of their target anthocyanin genes in all pigmented tissues and, consequently, to completely white flowers (Quattrocchio et al., 1993). The function of an2, on the other hand, seems highly redundant, because in an2− mutants, pigmentation of seeds, the flower stem, anthers, and the corolla tube is unaltered, whereas only pigmentation of the corolla limb is reduced. In an2− corolla limbs, some residual expression of target genes remains detectable (Quattrocchio et al., 1993); as a consequence, the corolla limb is patchy and pale colored rather than completely white. In a P. axillaris background, however, the formation of the pale petal color is prevented by the hf1-1 and Fl+ alleles, which in combination also reduce anthocyanin formation (Wiering and De Vlaming, 1984). In transient expression assays, we could not detect activity of the an2− alleles, even when they were expressed from the strong 35S promoter, indicating that they are null alleles. Therefore, it is possible that the residual expression of structural genes in an2− corolla limbs is controlled by another paralogous locus. We recently isolated a candidate AN2 paralog by yeast two-hybrid screens (A. Kroon and R. Koes, unpublished data), a finding that may help to solve this issue.

The Role of Transposons in the Generation of an2− Alleles

In natural isolates of P. axillaris, we found three different an2− alleles harboring either a nonsense mutation or a +4- or a −1-bp frameshift mutation. Our data strongly suggest that these were the mutations that inactivated an2 in nature. First, an2− Petunia spp lines still express an2 transcripts, indicating that cis-acting elements in the promoters of the corresponding an2 alleles are still working properly. The slightly lower abundance of an2 transcripts in an2− lines compared with An2+ lines may be due to the different genetic backgrounds of the Petunia spp lines and to reduced stability of an2 transcripts because of premature termination of translation. The latter phenomenon was also observed for transcripts with frameshifts of an11 (de Vetten et al., 1997), an1 (C. Spelt, F. Quattrocchio, J. Mol, and R. Koes, manuscript in preparation), and alf (Souer et al., 1998) and was reported by others as well (van Hoof and Green, 1996).

Because the +4- and the −1-bp frameshifts found in most an2− alleles are located on the same site, it is unlikely that they were generated by completely independent events. Closer inspection of the 4-bp insertion sequence shows that it has precisely the structure predicted for the insertion and subsequent excision of a transposon, strongly suggesting that the +4-bp and −1-bp an2− alleles were generated by two independent excisions of the same transposon (Figure 7). The petunia transposons dTph1 (Gerats et al., 1990), dTph2 (van Houwelingen et al., 1998), dTph3 (Kroon et al., 1994), dTph4 (Renckens et al., 1996; Alfenito et al., 1998), and dTph5 (C. Spelt and R. Koes, unpublished data) all belong to the Activator superfamily and generate an 8-bp target site duplication upon insertion. After excision, a footprint is left behind that consists of remains of the target site duplication (≤7 bp of each target site duplication; cf. Coen et al., 1986; Kunze et al., 1997) often separated by one or more inverted nucleotides of the target site duplication. Figure 7A shows that the −1- and +4-bp frameshift mutations can be explained, consistent with existing models for footprint formation, by two independent excisions of the same Activator-like transposon. However, in our analyses of footprints produced by dTph1 elements, we never found the relatively large deletions in the target site duplication inferred in Figure 7A, suggesting that these may be rare events (van Houwelingen et al., 1999).

More recently, two families of petunia transposons—Ps1 (Snowden and Napoli, 1998) and dTph6 (C. Spelt and R. Koes, unpublished data)—were discovered that belong to the so-called CACTA family. Ps1 and dTph6 also belong to the so-called CACTA family. These transposons generate a 3-bp

Figure 7. Model Showing How the Insertion and Excision of a Transposable Element May Have Created an2− Alleles with Two Different Frameshifts (−1 and +4 bp) at One Site.

(A) Model based on the insertion and excision of an Activator-like element, such as dTph1.

(B) Model based on the insertion and excision of a CACTA element, such as dTph6.

In (A) and (B), the nucleotides that are duplicated by the transposon insertion are underlined; nucleotides that have been lost during subsequent transposon excision and break-repair are shown as dots.
target site duplication upon insertion and produce footprints that consist of remains of the target site duplication (≤3 bp), often separated by small inversions of the target site duplication without missing nucleotides at the symmetry axis (cf. Coen et al., 1986; Kunze et al., 1997). Figure 7B shows that the −1-bp and +4-bp frameshift an2− alleles are fully consistent with the insertion and excision of a CACTA transposon. All of the above-mentioned transposons were identified in Petunia hybrida initially. However, DNA gel blot experiments showed that all P. axillaris and P. integrifolia accessions analyzed contained numerous copies of dTph1 (Gerats et al., 1990; Huits et al., 1994b), Ps1 (Snowden and Napoli, 1998), and dTph6 (F. Quattrocchio, C. Spelt, and R. Koes, unpublished data).

Taken together, these data suggest that in nature, the −1-bp and +4-bp frameshift in the an2− alleles of P. axillaris were most likely generated by a transposable element, presumably a member of the CACTA family, although the involvement of an Activator-like element cannot be excluded.

The Evolution of Flower Color in Petunia

The genetic separation of populations followed by mutation, selection, and genetic drift is thought to be the basis for the creation and divergence of species (Coyne, 1992). Traditionally, it is believed that this requires physical isolation, for instance, by geographical factors (allopatric speciation). However, speciation can in principle also occur between individuals that grow side by side (sympatric speciation). As flower shape and color attract specific pollinating animals, alteration of one or more of these traits may cause the attraction of different pollinators and genetic isolation (Vickery, 1992). For example, Mimulus spp mutations at only a few loci are sufficient to alter floral traits, such as shape, color, color pattern, and nectar yield, and switch from bee to hummingbird pollination (Bradshaw et al., 1995). Although this is consistent with a sympatric speciation scenario, it remains to be determined if Mimulus spp arose sympatrically. One problem with such sympatric speciation scenarios is to envision how the first mutation, for example, a new color, becomes fixed in the population, because it might initially be at a disadvantage as long as the flower retains its original shape.

The sequences of the an2 alleles provide a record of their history from which one can deduce how different flower colors arose in the genus Petunia in relation to speciation events. Figure 8 provides a model that summarizes and explains our findings. First, the presence of mutated an2− alleles in P. axillaris subsp indicates that the common ancestor of P. integrifolia and P. axillaris was colored flowering (An2+) and that the white (an2−) P. axillaris flowers arose by subsequent loss of an an2 function. This is in contrast to the evolution of kernel pigmentation in maize, which was established by gain-of-function mutations that altered the expression pattern of the regulatory c1 and r genes of its progenitor Te-osinte (Hanson et al., 1996). Second, we postulate that the ancestral petunia population split for unknown reasons (e.g., a geographical barrier or alterations in flower shape) into two genetically isolated groups that provided the foundation for today’s P. integrifolia and P. axillaris subsp.

The old polymorphisms in An2+ and an2− alleles are not spread randomly over the gene but rather are clustered in regions in which sequence alterations are relatively easily tolerated without loss of function, such as the intron and the 3’ end of the coding sequence (Figure 4). This indicates that the an2 gene was after the genetic separation, at least for some time, still under selection pressure in both Petunia spp populations. This implies that no inactivating mutations had occurred and that the flowers were still colored (An2+) at this stage of speciation. In the group that founded P. axillaris, at least two independent an2− alleles arose through the generation of a nonsense mutation and the insertion and excision of a transposon, respectively. Presumably, unknown circumstances (e.g., the insects visiting the flowers) favored white corolla limbs in this population, which explains why the an2− alleles became fixed and both the An2+ progenitor and An2− revertant alleles (in which transposon excision generated an active An2− allele) seem to have disappeared.

In this respect, it is interesting that insects (bees) are able to associate color with other characteristics such as scent and reward (Srinivasan et al., 1998). Because in P. hybrida, other anthocyanin genes, such as an1 (a regulatory gene) and an3 (encoding the enzyme flavanone 3β-hydroxylase), are more susceptible to transposon insertions than is an2 (van Houwelingen et al., 1998), it...
is likely that the an2- mutants were only a few out of many white-flowering petunia mutants that arose in nature. We assume that only the an2 alleles became fixed, because they only reduce anthocyanin synthesis in the petal limb, without additional effects on flavonoid accumulation. Mutation of any of the other known anthocyanin genes also reduces the synthesis of flavonoids in other tissues, such as the flower tube and the seed coat, and/or reduces the synthesis of other flavonoid compounds, such as flavonols (Quattrocchio et al., 1993; Huits et al., 1994a; van Houwelingen et al., 1998), which is apparently disadvantageous in a natural habitat.

Taken together, our data indicate that the flower color change caused by an2- mutations occurred after the genetic separation of both Petunia groups. It seems that the an2- mutation(s) was not the primary cause of speciation but rather a reinforcement mechanism (Coyne, 1992) that helped to complete the genetic separation.

METHODS

Plant Materials

The inbred lines of Petunia axillaris subsp. axillaris (S1 and S2) and subsp. parodi (S7 and S8) and P. integrifolia subsp. inflata (S6), subsp. violacea (S9 and S10), and subsp. integrifolia (S12) have been described previously (Wijsman, 1982). The P. axillaris accessions X504 and X508 (our numbering) were obtained around 1980 from the Botanical Garden of Liebig University (Giessen, Germany) and from R.N. Bowman (Goldsmith Seeds, Gilroy, CA), respectively. The an2-W82 allele was kept in the Amsterdam collection in the P. hybrida line W82 and is presumably identical to an2-n (Cornu, 1977).

Nucleic Acid Analysis

DNA extraction, DNA gel blot analysis, polymerase chain reaction (PCR) amplification, and sequence analyses were performed as previously described (Sour et al., 1995; de Vetten et al., 1997). The flanking sequence of the dTph1 element in an2-W82 (af41) was identified by established procedures (Sour et al., 1995; de Vetten et al., 1997) and used to screen a petal cDNA library and a genomic library of P. hybrida V26. an2 cDNAs from other lines were isolated by reverse transcription (RT)-PCR, and crucial regions were sequenced in at least two independently amplified products and in many cases by sequencing the corresponding region in PCR products amplified from genomic DNA. Analyses of DNA sequences were performed by using the program Geneworks 2.3 (IntelliGenetics, Mountain View, CA). Maximum parsimony analysis was performed with the program PAUP 3.1.1 (Smithsonian Institution, Washington, DC).

Quantification of mRNAs by RT-PCR analysis was performed as described previously (Quattrocchio et al., 1998).

Transient Expression Assays

an2 cDNA fragments were ligated between the cauliflower mosaic virus 35S promoter and nopaline synthase polyadenylation signal and delivered into an2-W115 mutant petal limbs by particle bombardment. AN2 activity was determined by measuring the induction of a dfrA-luc reporter gene and normalization for transformation efficiency using a 35S–β-glucuronidase gene, as described previously (de Vetten et al., 1997; Quattrocchio et al., 1998).

ACKNOWLEDGMENTS

We are grateful to Mark Rausher for discussions and review of the manuscript, Dik Roelofs for help with the parsimony analyses, and Jane Olsson for help in improving the manuscript. This work was supported by the Netherlands Organisation for Chemical research (S.O.N.), with financial aid from the Netherlands Organisation for the Advancement of Research (N.W.O.), and by the European Union BIOTECH program.

Received January 14, 1998; accepted May 19, 1999.

REFERENCES


Molecular Analysis of the *anthocyanin2* Gene of Petunia and Its Role in the Evolution of Flower Color
Francesca Quattrocchio, John Wing, Karel van der Woude, Erik Souer, Nick de Vetten, Joseph Mol and Ronald Koes
*Plant Cell* 1999;11:1433-1444
DOI 10.1105/tpc.11.8.1433

This information is current as of June 24, 2017

<table>
<thead>
<tr>
<th>References</th>
<th>This article cites 48 articles, 18 of which can be accessed free at: /content/11/8/1433.full.html#ref-list-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>eTOCs</td>
<td>Sign up for eTOCs at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>CiteTrack Alerts</td>
<td>Sign up for CiteTrack Alerts at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a></td>
</tr>
<tr>
<td>Subscription Information</td>
<td>Subscription Information for <em>The Plant Cell</em> and <em>Plant Physiology</em> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a></td>
</tr>
</tbody>
</table>

© American Society of Plant Biologists
ADVANCING THE SCIENCE OF PLANT BIOLOGY