12-2006

Convergence, Constraint and the Role of Gene Expression During Adaptive Radiation: Floral Anthocyanins in Aquilegia

Justen B. Whittall
Santa Clara University, jwhittall@scu.edu

Claudia Voelckel

Daniel J. Kliebenstein

Scott A. Hodges

Recommended Citation
Convergence, constraint and the role of gene expression during adaptive radiation: floral anthocyanins in Aquilegia. Whittall, Justen B.; Voelckel, Claudia; Kliebenstein, Daniel J.; et al. Molecular Ecology. December 2006, 15:14 pp. 4645-4657

Molecular Ecology (2006) **15**, 4645–4657 doi: 10.1111/j.1365-294X.2006.03114.x

Convergence, constraint and the role of gene expression during adaptive radiation: floral anthocyanins in *Aquilegia*

JUSTEN B. WHITTALL,*§ CLAUDIA VOELCKEL,* DANIEL J. KLIEBENSTEIN† and SCOTT A. HODGES*

*Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, CA 93106, USA †Department of Plant Sciences, University of California, Davis, One Shields Ave., Davis, CA 95616, USA

Abstract

Convergent phenotypes are testament to the role of natural selection in evolution. However, little is known about whether convergence in phenotype extends to convergence at the molecular level. We use the independent losses of floral anthocyanins in columbines (*Aquilegia*) to determine the degree of molecular convergence in gene expression across the anthocyanin biosynthetic pathway (ABP). Using a phylogeny of the North American *Aquilegia* clade, we inferred six independent losses of floral anthocyanins. Via semiquantitative reverse transcriptase–polymerase chain reaction (RT–PCR), we monitored developmental and tissue-specific variation in expression of the six major structural ABP loci in three *Aquilegia* species, two that produce anthocyanins (A+) and one that does not (A–). We then compared ABP expression in petals of old-bud and pre-anthesis flowers of 13 *Aquilegia* species, eight wild species and two horticultural lines representing seven independent A– lineages as well as three wild A+ species. We only found evidence of down-regulation of ABP loci in A– lineages and losses of expression were significantly more prevalent for genes late in the pathway. Independent contrast analysis indicates that changes in expression of dihydroflavonol reductase (DFR) and anthocyanidin synthase (ANS) are strongly phylogenetically correlated consistent with the multilocus targets of trans-regulatory elements in the ABP of other systems. Our findings strongly suggest that pleiotropy constrains the evolution of loss of floral anthocyanins to mutations affecting genes late in the ABP mostly through convergent changes in regulatory genes. These patterns support the hypothesis that rapid evolutionary change occurs largely through regulatory rather than structural mutations.

Keywords: anthocyanin biosynthetic pathway, *Aquilegia*, constraint, convergent evolution, gene expression

Received 24 April 2006; revision accepted 25 July 2006

Introduction

Changes in regulatory genes are predicted to predominate during morphological evolution, especially when it occurs rapidly (Britten & Davidson 1969; King & Wilson 1975; Barrier *et al*. 2001). To critically examine this prediction, we need to investigate the molecular events that caused the evolution of many traits in many species. However, on a smaller scale, we can test this hypothesis by studying cases where a single trait has evolved multiple times. Such convergent evolution is believed to be a consequence of parallel selection pressures exerted in similar environments (Simpson 1944; Stebbins 1974) and can occur across distantly related lineages (e.g. succulent stems evolved in at least nine plant families including Cactaceae and Euphorbiaceae, Mauseth 2004) or very closely related lineages (e.g. the repeated evolution of albinism in different populations of cavefish, Protas *et al*. 2006). The question of whether convergent phenotypes result from convergent changes in the same loci, or even in functionally similar genic regions (regulatory or coding), remains largely unexplored.

Justen B. Whittall and Claudia Voelckel contributed equally to this work.

Correspondence: Scott A. Hodges, Fax: (805) 893 4724; E-mail: hodges@lifesci.ucsb.edu.

Present address: Section of Evolution and Ecology, University of California, Davis, One Shields Ave., Davis, CA 95616, USA.

Theoretical analysis predicts that the probability of independent fixation of the same mutation is inversely related to the number of alternative beneficial mutations (Orr 2005). Because of the relatively few mutations that could lead to gain-of-function adaptations, the degree of their molecular convergence is predicted to be very high. For example, melanism in several animal species is caused by the same or similar amino acid substitutions in the melanocortin-1-receptor (Majerus & Mundy 2003) and insecticide resistance evolved through identical amino acid substitutions in several phylogenetically distinct insect lineages (Ffrench-Constant *et al*. 2000). In contrast, loss-of-function adaptations have been suggested to be less molecularly convergent (Orr 2005) because there are many potential mutations that could result in the loss of a phenotype. Not only are there many genes that could be potential targets, but, any number of mutations may render a gene nonfunctional. Furthermore, mutations to the *cis*-regulatory elements of a gene or its trans-regulatory factors could cause a loss-of-function through down-regulation of expression. However, there may be constraints that reduce the number of target genes for loss-of-function mutations if those genes serve essential functions elsewhere in the organism (Cresko *et al*. 2004; McCune & Carlson 2004).

One loss-of-function adaptation that is phenotypically and molecularly tractable is the loss of floral anthocyanins. Anthocyanins are the predominant floral pigments in angiosperms conferring red, blue and purple coloration to flowers. Flower colours function primarily as signals to pollinators and the evolution of new pollinator associations often are accompanied by shifts in flower colour (Fenster *et al*. 2004). However, flower colour may also be under selection pressures by other nonpollinator agents (reviewed in Strauss & Whittall in press). Anthocyanin biosynthesis is extremely well characterized across a diversity of angiosperm lineages at both the biochemical and molecular levels (Holton & Cornish 1995). The core anthocyanin biosynthetic pathway (ABP, Fig. 1) consists of only six structural genes conserved across angiosperms. However, intermediates in the ABP are used to produce many other flavonoid compounds that have been shown to be involved in UV-protection, edaphic tolerance, male fertility, herbivore and pathogen resistance and several other nonfloral functions (Koes *et al*. 1994; Steyn *et al*. 2002; Winkel-Shirley 2004). This diversity of functions increases the probability that loss-of-anthocyanin mutations could be constrained by pleiotropic effects.

Here, we study an example of repeated losses of floral anthocyanin production across several closely related North American *Aquilegia* species. The North American *Aquilegia* clade provides a unique system to test for molecular convergence in loss-of-function mutations and to examine the role of regulatory changes during rapid evolutionary change for several reasons. First, diversification

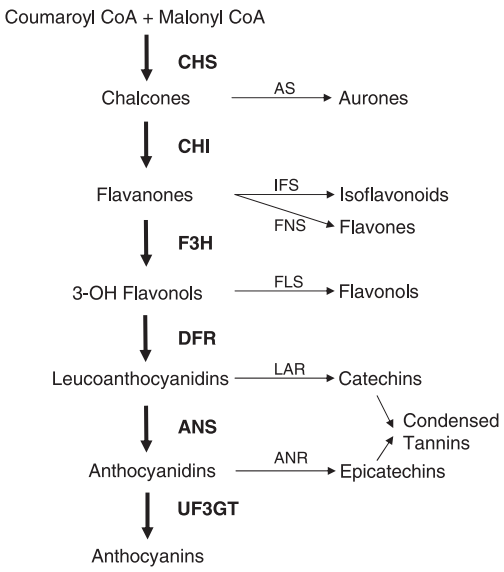


Fig. 1 The core anthocyanin biosynthetic pathway (ABP) consists of six genes: chalcone synthase (CHS), chalcone isomerase (CHI), flavanone-3-hydroxylase (F3H), dihydroflavonol reductase (DFR), anthocyanidin synthase (ANS), and UDP glucose: flavonoid 3-O-glucosyltransferase (UF3GT). Side-branch products are indicated with horizontal arrows along with their respective genes: aureusidin synthase (AS), isoflavone synthase (IFS), flavone synthase (FNS), flavonol synthase (FLS), leucoanthocyanidin reductase (LAR), and anthocyanidin reductase (ANR).

in *Aquilegia* has happened very recently (Hodges & Arnold 1994; Whittall *et al*. 2006), making the causal mutations less likely to be obscured by subsequent mutations. Second, the loss of floral anthocyanin production causing a shift to yellow or white flower colours (depending on the presence of carotenoids) is likely adaptive as it is generally correlated with hawkmoth pollination (Fenster *et al*. 2004) and many species of *Aquilegia* show this association (Miller 1981, 1985; Grant 1992; Hodges *et al*. 2002, 2004). In fact, in choice experiments hawkmoths prefer paler flowers (Hodges *et al*. 2004) and in populations that are polymorphic for flower colour, white flowers had a fitness advantage when hawkmoths were the major pollinator (Miller 1981). Third, we have reconstructed a phylogeny for the North American *Aquilegia* clade (Whittall 2005) allowing us to determine the evolutionary history of anthocyanin production through phylogenetic character mapping.

Using the *Aquilegia* phylogeny, we estimate the number of independent losses of floral anthocyanins. Then, to determine the role of gene regulation underlying this morphological convergence, we compare expression of all six structural ABP loci across 11 *Aquilegia* species and two horticultural lines including both those that produce anthocyanins (A+) and those that do not (A–). In addition, we verify the functionality of the potentially constrained

portions of the ABP through biochemical detection of flavonoid intermediates in flowers of several A– and A+ taxa.

Materials and methods

Ancestral state reconstructions

North American *Aquilegia* species have been found to either produce delphinidin-based (blue/purple, *A. brevistyla* and *A. coerulea*) or a combination of pelargonidin and cyanidin-based (red and blue, respectively, *A. formosa*, *A. canadensis*) forms of anthocyanins (Taylor 1984). Floral anthocyanins were coded as present or absent in all North American *Aquilegia* species based on one or more of the following criteria: average reflectance curves (Hodges *et al*. 2002), degree of anthocyanin pigmentation in the vacuole under compound microscopy (×100) and/or field observations of several populations per species. Although flower colours are nearly always homogeneous within an individual, some *Aquilegia* spp. have variable levels of anthocyanin pigmentation within and among populations (*A. pubescens*, *A. coerulea* var. *coerulea*, *A. coerulea* var. *ochroleuca*, *A. coerulea* var. *alpina*, *A.* sp. nov. and *A. scopulorum*; Grant 1952; Miller 1981; Whittemore 1997) and were coded as polymorphic. The presence, absence and polymorphisms of floral anthocyanins were mapped onto an amplified fragment length polymorphism (AFLP) phylogeny (Whittall 2005). Proportional likelihoods for ancestral state reconstructions and tree-wide rates of anthocyanin gain and loss were estimated using the maximum-likelihood (ML) option in BAYESMULTISTATE (Pagel *et al*. 2004). For polymorphic taxa, likelihoods were calculated over both states assuming equal probability of each state providing a conservative estimate of ancestral states and rates of character evolution (Pagel *et al*. 2004). A model of character evolution with asymmetric rates of gains and losses was tested against a symmetrical one-rate model using the RESTRICT OPTION in BAYESMULTISTATE. Statistical significance was determined with a likelihood-ratio test where the degree of freedom is equal to the difference in the number of free parameters in two nested models (d.f. = 1). The degree of asymmetry in the rates of gains and losses was tested by restricting the gain rate to zero and comparing the resulting likelihood score to the unrestricted model in a likelihood-ratio test (d.f. = 1). The robustness of these results to different character state assignments for the polymorphic taxa was tested using alternative character codings for the polymorphic taxa.

Plant material

Floral tissue was collected from 13 *Aquilegia* species: 11 North American species (*A. laramiensis*, *A. canadensis*, *A. longissima*, *A. pinetorum*, *A. chrysantha*, *A. micrantha*, *A. barnebyi*, *A. coerulea* var. *coerulea*, *A. pubescens*, *A. flavescens*, *A. formosa*), one horticultural derivative of *A. coerulea* (*A. coerulea* 'Origami white') and one horticultural derivative of Japanese *A. flabellata* (*A. flabellata* var. *pumila f. alba*). The two horticultural lines were included to represent independent origins of A– from the *A. coerulea* lineage and an old world lineage, respectively. Specimens were greenhouse-grown transplants from wild populations or grown from wild-collected seed except for *A. laramiensis* and the two horticultural lines which were obtained from nurseries in Medford, OR and Gilroy, CA. For *A. canadensis* (A+), sepals and petals were harvested from flowers of five developmental stages (green bud, young bud, old bud, pre-anthesis, and post-anthesis, see Fig. S1a, Supplementary material) as well as pre-anthesis anthers from three individuals. Anthocyanins are visually obvious from the young bud stage through post-anthesis stage in sepals and petals. For both, *A. formosa* and *A. chrysantha*, petals from the above five developmental stages were sampled whereas for each of the remaining species petals from only two stages (old bud, pre-anthesis) were included in the analysis. The blade and the nectary were removed from all petals. Tissues were snap-frozen under liquid nitrogen and stored at –80 °C.

RNA isolation, cDNA synthesis and DNA isolation

Total RNA was isolated from 50 to 100 mg of ground tissue using the Mini-to-Midi RNA extraction kit (Invitrogen). Before complementary DNA (cDNA) synthesis, RNA samples were treated with DNAse I (Invitrogen) and cDNA was synthesized using oligo dT primers and the Supercript III First Strand Synthesis Kit (Invitrogen). DNA of species that were lacking expression of a single or multiple ABP loci was isolated from 100-mg fresh leaf tissue using the Plant Mini Kit (QIAGEN).

ABP loci and primer design

We chose to study the core ABP genes (Fig. 1). There are other genes that utilize intermediates in this pathway. In particular, two genes, F3′H and F3′5′H affect the class of anthocyanins produced, cyanidins and delphinidins, respectively, rather than pelargonidins, which are produced if neither gene is active. *Aquilegia formosa* and *A. canadensis* produce both pelargonidins and cyanidins (Taylor 1984) indicating that the loss of F3′H would simply cause only pelargonidins to be produced. *Aquilegia brevistyla* and *A. coerulea* produce delphinidins (Taylor 1984). The loss of F3′5′H in these species would cause white flowers only if any of the downstream genes were specific for their products, otherwise simply a change in the class of anthocyanins would occur. This transition is rare and therefore, for this study, we chose not to investigate the expression patterns of either F3′H or F3′5′H.

Degenerate primers were developed using CODEHOP (Rose *et al*. 1998) for all six ABP loci. Because *Aquilegia* is a basal eudicot, phylogenetically distant from model systems, we used amino acid alignments of 10–15 full-length cDNA sequences from both monocots and eudicots to identify regions of conserved amino acids for primer design (Table S1, Supplementary material). Our aim was to produce primers that (i) would likely amplify all paralogous loci if they were expressed; and (ii) would amplify across multiple species even if there was some nucleotide variation at the primer sites. Primer sequences were determined based on the codon frequency table of *Arabidopsis thaliana* and were adjusted to match any conserved nucleotides within the alignments.

ABP loci were isolated from floral cDNA of two A+ species, *A. formosa* [chalcone syntase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), anthocyanidin synthase (DFR), anthocyanidin synthase (ANS)] and *A. flabellata* [UDP-glucose flavonoid 3-O-glucosyltransferase (UF3GT)] via reverse transcriptase–polymerase chain reaction (RT–PCR) [final concentrations of PCR ingredients: 0.25 mM dNTP, 2.5 mM $MgCl_2$, 1× PCR buffer B, 1 μM forward primer, 1 μM reverse primer, 1 U *Taq* Polymerase (Fisher)]. Thermal cycling conditions were 92 °C for 5 min, followed by 35 cycles of 92 °C for 15 s, annealing for 45 s and 72 °C for 1 min 30 s followed by 72 °C for 10 min. Optimal annealing temperatures to obtain a single band were determined empirically. For primer sequences and annealing temperatures see Table 1. RT–PCR products were gel purified (QIAGEN) and cloned (TOPO-TA cloning kit, Invitrogen). Colonies containing inserts of the predicted size were sequenced using primer-labeled simultaneous bidirectional sequencing (SBS) on LI-COR 4200 DNA sequencers (LI-COR Biosciences).

RT–PCR optimization and experimental design

A template dilution series from *A. canadensis* cDNA was amplified and the number of PCR cycles that resulted in agarose-gel band intensities that enabled us to distinguish among different amounts of starting material was used for all subsequent PCRs (Table 1).

In order to determine individual, developmental and tissue-specific expression patterns, we compared expression from three individuals, five floral developmental stages and three floral tissues of *A. canadensis*. We then examined expression patterns among the same five floral stages of *A. formosa* (A+) and *A. chrysantha* (A–). In a third experiment, we compared the expression of the six ABP loci in late floral buds and pre-anthesis flowers across 13 *Aquilegia* species. All reactions were carried out using 20 ng of reverse-transcribed RNA with the above PCR conditions and each amplification included samples of nonreverse-transcribed RNA to control for amplification of contaminating DNA. Bands were quantified using NIH IMAGE (http://rsb.info.nih.gov/nih-image/) and ACTIN1 levels were used to standardize ABP gene expression. An arbitrary threshold of 50% of maximum expression was chosen to designate genes as being down regulated.

To test for correlated changes in expression in a phylogenetic framework, we used independent contrasts based on the AFLP phylogeny (Whittall 2005). Independent contrasts is a commonly used comparative technique assuming a Brownian motion model of evolution to test for evolutionary correlations of continuous traits after explicitly accounting for phylogenetic relatedness (Felsenstein 1985; Garland *et al*. 2005). Initial analyses were carried out using the phylogeny pruned to the 11 *Aquilegia* species sampled for gene expression without the horticultural samples. Subsequent analyses included the horticultural samples by associating their expression data with the phylogenetic position of their stock parental lineages. For both analyses, we used the PHYLOGENETIC DIVERSITY ANALYSIS PROGRAMS (PDAP; Midford *et al*. 2005) module in MESQUITE version 1.06 (Maddison & Maddison 2005) to examine independent contrasts for all pairwise comparisons of ACTIN-normalized gene expression values for the six ABP genes and a Bonferroni correction to determine significance at $\alpha = 0.05$. Results using the AFLP branch lengths were compared with three alternative branch-length assignments: all branch lengths = 1, Grafen branch lengths, and an ultrametric tree in MESQUITE version 1.06 (Maddison & Maddison 2005).

In order to rule out priming-site mismatch as an explanation for low or absent RT–PCR product, ABP loci were amplified from genomic DNA using the same PCR conditions (except for a 2.50-min elongation step to allow for potentially longer DNA fragments to amplify) for selected A- species and several A+ species. Results were compared on 1.5% TBE agarose gels.

HPLC survey of flavonoid intermediates

Two classes of flavonoid intermediates — flavones and flavonols — were quantified from 100 mg of fresh sepal tissue via high-performance liquid chromatography (HPLC) analysis (Merken *et al*. 2001) from *A. coerulea* var. *ochroleuca* ($N = 6$), *A. chrysantha* ($N = 6$), *A. flavescens* ($N = 4$), *A.* sp. nov. ($N = 2$), *A. micrantha* ($N = 2$), *A. pubescens* ($N = 2$), *A. coerulea* var. *coerulea* ($N = 5$), *A. canadensis* ($N = 3$) and *A. scopulorum* A + ($N = 2$). The tissue was homogenized in a 90:10 methanol:water solution. Flavonoids were extracted by shaking at 8000 *g* for 4 h in the dark. After a 10-min centrifugation at 8000 *g*, 5 μL of the supernatant were injected into an Agilent HPLC-DAD machine and separated in a reverse phase C18 column with a flow rate of 1 mL/min. A solvent gradient consisting of acetonitrile and 0.1% trifluoroacetic acid ran for 1 h (Merken & Beecher 2000). Compounds were identified based on reference spectra

Table 1 Summary of PCR primers and characteristics of their amplification products in *Aquilegia*. For each anthocyanin biosynthetic pathway locus (for locus abbreviations, see Fig. 1), the names and sequences of both primers (degenerate 3′ core in lowercase letters, nondegenerate 5′-clamp in capital letters), the number of cDNA clones sequenced, the length of each PCR product and the accession number and taxon for the most similar GenBank sequence from a translated blast search (TBLASTX) are given. The estimated percentage of the open reading frame (ORF) for each *Aquilegia* cDNA was determined using the average length of 10–15 angiosperm ORFs (see Table S1, Supplementary material). Two primer sets were designed for DFR, one partially degenerate set and one *Aquilegia*-specific set

Locus	Primer name: forward reverse	Primer sequence	No. clones sequenced	PCR product length	Closest hit in TBLASTX search: taxon Genbank	*E*-value	Estimated % of ORF	Ann. temp. (°C)	RT–PCR cycles
CHS	AA113	F: AGAGGCTGCCCAGAAGGCCATCaargartgggg	13	756	*Petrosilenum crispum* V01538	$8E^{-139}$	63	60	26
	AA377	R: TTCAACAGTGAGTCCAGGCccraanccraa							
CHI	AA49	F: GAGGACTGGAAATTCAAGGTAAGTTCrtnaarttyac	11	497	*Citrus sinensis* AB011794	$5E^{-62}$	71	55	31
	AA208	R: GCTTTGCTGCCGGTGACACAccrtkytyncc							
F3H	AA26	F: ACTCGAACTTCGTGCGggangargayga	21	789	*Anethum graveolens* AY817679	$3E^{-161}$	72	55	31
	AA299	R: GTCGCCTCTGGTGCCggrttytgraa							
DFR	AA41	F: CTACACTGTTCGTGCCACTgtnmgngaycc	15	808	*Vitis vinifera* Y11749	$2E^{-147}$	67	55	
	AA321	R: CAGTCGATGGCTCCCTTGwmcatntcytc							
DFR	F64	F: ATGGAAAGCAGACCTTGTCG	NA	437	NA	NA	40	60	26
	R539	R: GTGCCTCACTACGGGTGATT							
ANS	AA88	F: GCTGCCATGGACTGgggngtnatgca	12	564	*Citrus sinensis* AY500593	$2E^{-73}$	52	55	32
	AA293	R: CCATTGCTCAGAATCTCAAGTGTGtcnccnayrtg							
UF3GT	AA28	F: GCTcccaTGTTGCTGTtstngcnttycc	13	994	*Aralia cordata* AB103471	$1E^{-102}$	72	55	32
	AA403	R: CTCTCCAAcACcGAGTTCcanccrsmrtg							
ACTIN	F	F: GATGGATCCTCCAATCCAGACACTGTA	NA	NA	NA	NA	NA	60	26
	R	R: GTATTGTGTTGGACTCTGGTGATGGTGT							

(Merken & Beecher 2000) and peak areas were measured at 330 nm. If a class of flavonoids was represented by multiple peaks, the peak areas were summed. Nonphylogenetic comparisons between flavonoid intermediates of A+ and A– species were evaluated by ANOVA and Fisher's protected least significance difference (PLSD) post hoc tests when sample sizes permitted. We also used independent contrasts to test for correlated changes in the concentration of flavonoid intermediates with the loss of anthocyanins in the PDAP module in MESQUITE as described above for the expression data.

Results

Estimating the number of independent losses of anthocyanins

A Bayesian consensus phylogeny of the North American *Aquilegia* clade pruned to species was used to determine the number of convergent losses of floral anthocyanins (Whittall 2005). Maximum-likelihood ancestral state mapping of floral anthocyanin production indicates five fixed losses (*Aquilegia flavescens*, the *Aquilegia micrantha* lineage, the *Aquilegia chrysantha* lineage, *Aquilegia longissima* (TX), *Aquilegia laramiensis*) and at least two additional independent losses in polymorphic taxa (*Aquilegia pubescens*, the *Aquilegia coerulea* lineage, Fig. 2). No gains are indicated in the ancestral state reconstructions. Note that taxa lacking anthocyanins have either yellow or white flowers depending on the presence or absence of carotenoids, which are long-chain carbon compounds from an unrelated biochemical pathway (Hirschberg *et al*. 1997). The ancestral state of polymorphic *Aquilegia scopulorum* is uncertain (likelihood proportions are 0.8 A–, 0.2 A+) and thus the taxon was not counted as an additional loss. In addition, the unexpected polyphyly of *A. longissima* (TX) suggests an independent origin of anthocyanin loss within this taxon. We do not count the loss in *A. longissima* (TX) until additional evidence confirms its phylogenetic position. Therefore, we conservatively estimate six independent losses of floral anthocyanins in the North American *Aquilegia* clade (*A. flavescens*, *A. pubescens*, the *A. micrantha* lineage, the *A. coerulea* lineage, the *A. chrysantha* lineage and *A. laramiensis*).

The transitions between A+ and A– phenotypes are better explained by an asymmetric model of anthocyanin evolution, where the rate of anthocyanin loss is significantly higher than the rate of anthocyanin gain (loss rate = 5.21, gain rate = 0.01; LR test for asymmetry $P = 0.0386$). Based on the ancestral state reconstructions, no reversals from A- to A+ phenotypes were inferred, but the estimated rate of gain was significantly greater than zero ($P < 0.01$). The maximum-likelihood estimation for polymorphic taxa weighs both states equally. This allows polymorphic nodes

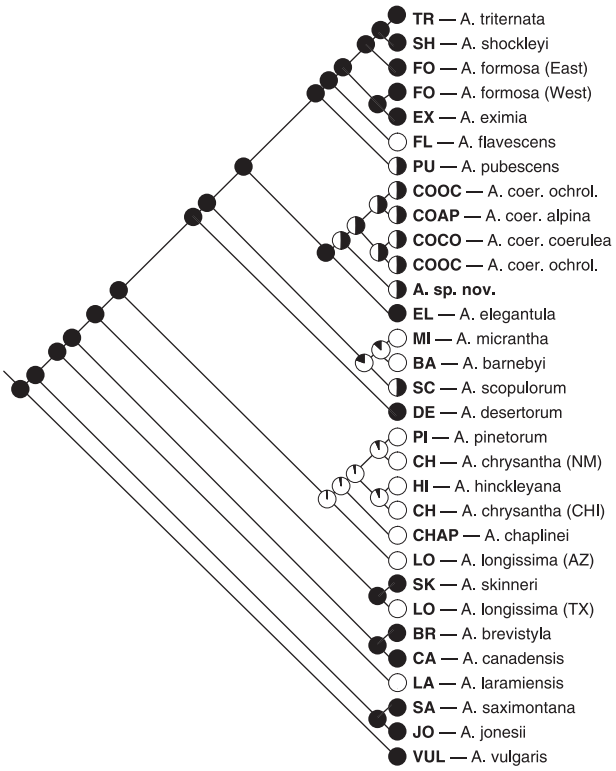


Fig. 2 Maximum-likelihood ancestral state reconstructions of floral anthocyanin production in *Aquilegia* based on the Bayesian AFLP phylogeny under a two rate model of evolution. Pie diagrams represent the likelihood proportion of each ancestral state: with anthocyanins (A+), filled circles; without anthocyanins (A–), empty circles. Taxa coded as polymorphic are indicated with half-filled circles.

arising from A+ ancestors, like that leading to *A. pubescens*, to either be considered losses (when calculated as A–) or no change (when calculated as A+). The strongly asymmetric rates are then driven by the remaining nonpolymorphic A– taxa, which represent mostly unambiguous losses [*A. laramiensis*, the *A. chrysantha* lineage, the *A. micrantha* lineage, *A. flavescens* and *A. longissima* (TX)]. Several alternative character state assignments among the varieties of the *A. coerulea* lineage and the exclusion of *A. longissima* (TX) produce nearly significant asymmetric rates of gain and loss ($P = 0.052$).

Isolation of ABP structural loci and PCR optimization

RT–PCR of *Aquilegia formosa* and *Aquilegia flabellata* cDNA from petal and sepal tissue with degenerate primers generated mostly single products in the expected size range for all six ABP loci (Table 1). We sequenced between 11 and 21 (mean = 14.2) clones per locus. Thus, if two paralogous loci were expressed in equal proportions and

had equal cloning efficiencies, there was a greater than 95% chance to detect both loci. We did not observe nucleotide variation beyond that predicted by *Taq* error (Beltran *et al*. 2002), indicating that only single loci were expressed for each gene in petal and sepal tissue. We used translated blast searches (TBLASTX) to identify the most similar characterized genes for each *Aquilegia* cDNA sequence. In every case, the most similar GenBank sequence was identified as the target locus and *E*-values were very low (5E^{-62} or lower, Table 1). The conserved regions, identified in the angiosperm alignment used in primer design, were also found in each *Aquilegia* sequence (Table S1). By designing degenerate primers to highly conserved regions, we increased the probability of monitoring the expression of paralogous gene copies that may be expressed in those *Aquilegia* species that were not initially used to clone all loci.

The six ABP loci were optimized for annealing temperature (Table 1). Because nonspecific amplification products were produced with the degenerate primers for DFR, we designed gene-specific primers for this locus (Table 1). After these optimizations, all primers produced strong, single RT–PCR amplification products in several phylogenetically distant A+ *Aquilegia* species. We determined the optimal number of PCR cycles to distinguish differences in the amount of starting material for each locus (Table 1) by amplifying from a dilution series of *Aquilegia canadensis* cDNA template.

Individual, developmental and tissue-specific variation in ABP gene expression

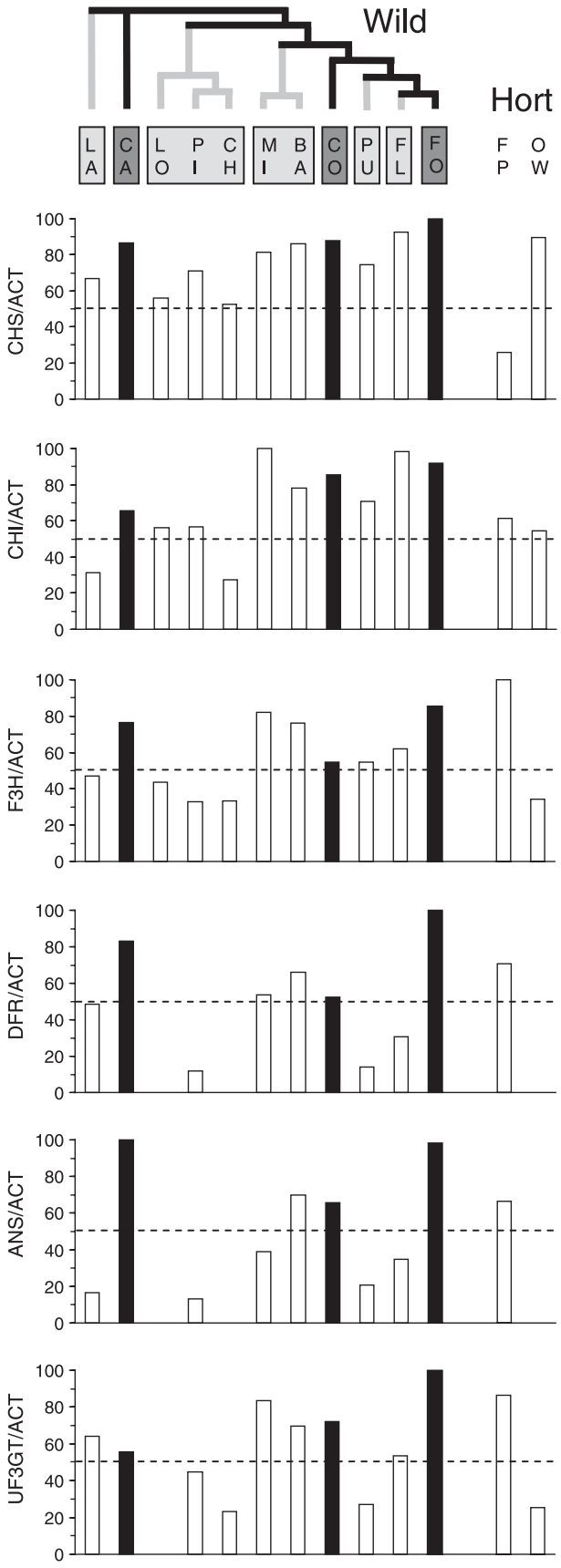
We found all six structural ABP loci to be expressed in a development-specific manner in each of the three individuals of the A+ species, *A. canadensis* (Fig. S1a, b, supplementary material). In petals, only CHI was expressed invariably from green buds to post-anthesis flowers; F3H expression peaked in old buds and then declined; the expression of CHS, DFR, ANS and UF3GT showed an initial increase and then remained fairly constant in later floral stages. The patterns for CHS, CHI and F3H in petals were mirrored in sepals. For DFR, ANS and UF3GT, however, expression patterns in sepals were somewhat different from that found in petals with expression declining in later developmental stages. Across the five developmental stages, in both tissues, actin-normalized expression levels of DFR and ANS were highly correlated (petals: $y = 139.8 + 97.8 \times \log(\text{ANS})$, $r = 0.954$, $F_{1,13} = 132.5$, $P < 0.0001$; sepals: $y = 239.4 + 69.8 \times \log(\text{ANS})$, $r = 0.894$, $F_{1,13} = 51.7$, $P < 0.0001$). These correlations were the highest found for expression levels between any pair of loci (data not shown). *Aquilegia canadensis* anthers, which do not produce anthocyanins, expressed CHS and CHI at levels comparable to those of petals and sepals, while F3H and UF3GT showed a severely reduced expression, and DFR and ANS were not expressed at all (Fig. S1a, b, supplementary material).

For most loci, except ANS, we observed greater individual variation in sepals than in petals, and we therefore conducted further experiments with petal tissue. The highest expression levels were found in old buds and pre-anthesis flowers for nearly all genes, which rendered these stages preferable for a cross-species comparison (Fig. S1a, b).

To examine whether other A+ species have similar developmental expression patterns as found in *A. canadensis* and to evaluate transcription in an A– species, we determined expression patterns from petal tissue in one individual of *A. formosa* (A+) and one individual of *A. chrysantha* (A–). Expression patterns for CHS and CHI in *A. formosa* were extremely similar to *A. canadensis*. Peak expression of F3H in *A. formosa* occurred in the old bud rather than the pre-anthesis stage. Expression patterns between *A. formosa* and *A. canadensis* were very similar for DFR, ANS and UF3GT although absolute levels were lower in *A. formosa*, especially for ANS and UF3GT. Similar to *A. canadensis*, the expression levels of DFR and ANS were highly correlated in *A. formosa* ($y = 172.3 + 0.91 \times \log(\text{ANS})$, $r = 0.893$, $F_{1,3} = 24.9$, $P = 0.0154$). Like in *A. canadensis* this correlation was the highest found for expression levels between any pair of loci (data not shown). The patterns found for *A. chrysantha* were different from those in the A+ species: CHS expression abruptly declined in post-anthesis flowers; CHI expression steadily decreased across floral developmental stages; F3H expression peaked earlier than in both A+ species; DFR and ANS transcripts were not detectable except for very weak expression in green buds and post-anthesis flowers, respectively; and UF3GT was only weakly expressed in old buds and pre-anthesis flowers (Fig. S2a, b, Supplementary material). Because the old bud and pre-anthesis flowers in these three species had the peak expression for nearly every gene, we used these same developmental stages in a larger cross-species comparison.

ABP gene expression patterns in thirteen Aquilegia *samples*

We monitored ABP gene expression in old buds and pre-anthesis flowers in three A+ species, eight A– species and two A– horticultural lines. Actin-normalized average expression levels, scaled as a percentage of the maximal level measured, are depicted in Fig. 3. The three A+ species (*A. canadensis*, *A. coerulea* var. *coerulea*, and *A. formosa*) expressed all six ABP loci at 50% of the maximal level or more. Because we sampled multiple species from some A- lineages, we considered an entire lineage as being down regulated when all species showed less than 50% of the maximal expression level measured. With this criterion, 18 genes had reduced expression levels across the seven



A– lineages. Of the seven A– lineages only one, *A. micrantha/barnebyi*, did not have any gene consistently down regulated. For a representative gel image, see Fig. S3, Supplementary material. Similar results were obtained by only considering the maximal expression values measured across the two developmental stages for each species (data not shown). In this case, 14 genes were still classified as having reduced expression levels. Only F3H in all spp. of the *chrysantha* clade, ANS in *A. flavescens*, and UF3GT in *Aquilegia pinetorum* were no longer considered down regulated.

To test whether down-regulation in A– lineages is randomly distributed in the ABP we used 2×2 conditional binomial exact tests (Rice 1988). We grouped the genes into 'early' vs. 'late' parts of the pathway. In some systems, F3H is considered an early gene and in others a late gene (Mol *et al*. 1998) and thus we conducted each test twice, once for each grouping. First, we counted the number of genes down regulated and not down regulated in each part of the ABP across all lineages. This test assumes that the expression pattern of each gene is independent (i.e. variation in regulation is due to *cis*-regulatory elements). We found significantly more genes down regulated in the 'late' part of the pathway whether F3H was counted as an 'early' or as a 'late' gene ($P = 0.0172$ and $P = 0.0096$, respectively). However, because expression of multiple ABP genes can be controlled by a common trans-regulator (Mol *et al*. 1998), the expression patterns of individual genes may not be independent. We therefore applied a very conservative test by only considering whether one or more genes were down-regulated in each part of the pathway. In this case the sample size was only seven (each A– lineage considered independently). Regardless of whether F3H is counted as an early or late gene, five of the seven A– lineages have multiple genes down regulated in the late part of the pathway as compared to only one A– lineage in the early part (*A. laramiensis* when F3H is considered an early gene). Neither test, however, was significant (F3H considered early or late, $P = 0.77$, 0.17, respectively).

To further investigate if trans-regulatory changes may be responsible for these down-regulations, we examined the actin-normalized expression levels. We observed that when a lineage is down regulated for one gene, it is always

Fig. 3 Cross-species comparison of gene expression in the six core ABP loci for 13 *Aquilegia* samples. Phylogenetic relationships among the 11 wild species and two A– horticultural varieties are depicted in the top panel with the dark and light lines indicating A+ and A– species, respectively (for species abbreviations see Fig. 2). Expression results are scaled to the highest expression ratio obtained for each locus. Solid and empty bars represent A+ and A– species, respectively; all bars represent the average values between old buds and pre-anthesis flowers of a single individual. Dashed lines denote 50% expression of the highest expression level measured.

down regulated for at least one other gene. In particular, the expression levels of DFR and ANS were highly correlated across species and stages ($y = 0.004 + 0.51 \times$ ANS, $r = 0.873$, $F_{1,30} = 95.8$, $P < 0.0001$). Like the correlations of expression levels among developmental stages of *A. canadensis* and *A. formosa* this correlation was the highest between any pair of loci (data not shown). In addition, using independent contrasts based on the underlying phylogenetic relationships of the samples, there is only one very strong correlation between expression levels for pairs of genes, that between DFR and ANS ($P = 4.76 \times 10^{-5}$). In all 12 contrasts, whenever DFR expression changes, ANS expression changes in the same direction. These results are robust to the exclusion of the horticultural lineages and to several alternative branch-length assignments (Table S2, Supplementary material).

In summary, a lack of floral anthocyanins correlates very strongly with reduced expression or a lack of expression of the late ABP loci (starting from F3H), particularly of DFR and ANS. In one lineage (*A. micrantha*/*A. barnebyi*), however, there was no correlation between ABP gene expression and a lack of floral anthocyanins.

ABP gene amplification from genomic DNA of selected A– species

To test whether variation at the priming sites may cause a failure of RT–PCR, we amplified those loci that had reduced or undetectable RT–PCR amplification products, using genomic DNA template for the respective A– species as well as several A+ species. We were able to amplify genomic F3H and ANS for all A– species with reduced or undetectable F3H and ANS expression as well as several related A+ species (data not shown). Amplification products for these genes from A– species were of similar size and intensity as the products from A+ species indicating that the degenerate primers were equally effective at PCR amplification across all species. We did not succeed in amplifying DFR or UF3GT from genomic DNA for any species (neither A+ nor A–), suggesting the presence of an intron spanning a priming site or very large introns between the primers. For example, sequence comparisons revealed that the specific reverse primer for DFR spans the location for an intron found in *Arabidopsis* (data not shown).

HPLC analysis of flavonoid intermediates

We measured flavone and flavonol production in three A+ species and a subset of six A– species to test for the presence of CHI activity and F3H activity, respectively. Despite species-specific differences in flavone and flavonol levels, we found all A– species capable of producing both classes of flavonoids within the range of those produced in A+ samples (Fig. 4). Independent contrasts revealed no significant correlation between flavonoid concentration and loss of anthocyanins (2-tailed sign test-flavones $P = 1.0$; flavonols $P = 0.73$). The presence of these two anthocyanin intermediates in A- species at levels comparable to those in A+ species indicates functional enzymes exist in flowers for the early ABP steps in our sample of A– species.

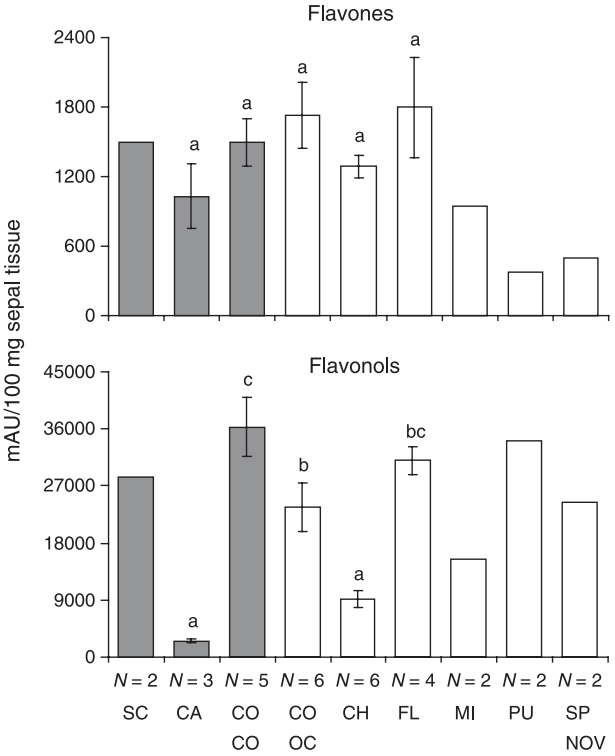


Fig. 4 Average peak areas of flavones and flavonols from nine *Aquilegia* species (shaded bars, A+ white bars, A–) measured by HPLC. For species abbreviations see Fig. 2. Standard errors are given only for species with at least three independent measurements. Taxa with lower case letters have significant differences in their flavone/flavonol production. Note similar levels of flavones in A– and A+ species ($F = 1.22$, $P = 0.3354$) but significant differences of flavonols among both A+ and A– species ($F = 16.803$, $P < 0.001$, Fisher's PLSD post-hoc test $P < 0.05$).

Discussion

In this study, we examined gene expression for six independent losses of floral anthocyanins found in the North American *Aquilegia* radiation and one loss that was derived from an Asian species, *Aquilegia flabellata*. By comparing many independent origins among closely related lineages, we greatly increased our power to detect general patterns and constraints during the evolution of white or yellow flowers. Our results suggest that the independent losses of floral anthocyanins in *Aquilegia*

are largely convergent at the expression level with most losses of expression concentrated in the late ABP genes, particularly DFR and ANS. Thus, loss-of-function phenotypes may be molecularly convergent, despite the seemingly large number of mutational targets — six genes in this case — that could potentially cause the phenotype.

A corollary of this conclusion is that strong negative pleiotropic effects preclude selection for mutations early in the pathway. CHS and CHI play an essential role in the first two dedicated steps of flavonoid metabolism and are known to have pleiotropic effects in other species. For example, white-flowered *Ipomoea purpurea* plants, which are homozygous for a structural mutation in the CHS gene, have reduced fertility, lower seed germination rates, decreased seedling survival, and an intolerance to heat stress (Coberly & Rausher 2003; Fehr & Rausher 2004). The decreased fitness associated with these mutants likely prevents the allele from spreading in natural populations. Similar selective forces may prevent the loss of expression of CHS and CHI in floral tissue of A– *Aquilegia* species. Moreover, the expression of CHS and CHI in stamens of *Aquilegia canadensis* (A+) is consistent with the role of flavonoid function in tissues that do not produce anthocyanins.

Not only do our data suggest a common region of the ABP pathway for loss-of-function mutations to be selected, but they also indicate a common specific target, trans-regulatory genes. Five A– lineages exhibited reduction in the expression of multiple genes late in the pathway. In particular, expression of DFR and ANS was reduced in each of these cases. These two genes appear to be co-regulated as their expression levels are highly correlated across developmental stages and tissues within species as well as across species. This hypothesis is consistent with the multilocus targets of regulatory elements that have been identified in other species (bHLH, MYB and WD40 proteins), which together regulate either the entire ABP or control expression of two or more genes downstream of CHI or F3H (Mol *et al*. 1998). Thus, mutations in a common regulator of DFR and ANS would explain their co-reduction across multiple lineages. Such mutations have been found in other systems including *Petunia* and *Ipomoea* where anthocyanin polymorphisms are caused by different alleles at trans-acting regulatory loci (Epperson & Clegg 1988; Quattrocchio *et al*. 1999; Chang *et al*. 2005). Also, in *Ipomoea*, Durbin *et al*. (2003) examining three ABP genes in several A– species, found reduced expression at multiple loci in most cases, and concluded that regulatory genes were the most common cause of A– phenotypes. Thus, down-regulation through mutations in trans-regulators may be a common mode for the evolution of anthocyanin-less flowers in many angiosperms.

An alternative explanation for most A– phenotypes is that one mutation caused the initial loss of anthocyanins and then subsequent mutations have caused lowered gene expression. For instance, a structural mutation in any ABP enzyme could block anthocyanin production initially. Then, either drift or selection could lead to reduced expression of the nonessential proteins. Such a scenario was recently suggested by Zufall & Rausher (2004) who investigated a shift from blue to red colour in *Ipomoea quamoclit*. They discovered two changes in the ABP that could have lead to the red phenotype: loss of F3′H expression — the enzyme that produces the precursor for the blue cyanidin-3-glucoside — and a significantly reduced ability of DFR to metabolize the product of F3′H, dihydroquercetin. One of these mutations may have originally led to the red phenotype and the second mutation may have arisen as a consequence of this modification (Zufall & Rausher 2004). Because we never detected either DFR or ANS to be down regulated independently, a multiple mutation hypothesis would imply that a mutation to one gene (or its unique trans-regulator) precipitates strong selection for down-regulation of the other gene. While we cannot rule out such a scenario, mutation in a common trans-regulator is a more parsimonious explanation. In order to test if a reduction in expression in A– *Aquilegia* species is preceded or succeeded by structural mutations that abolish enzymatic function, both sequence information and functional assays of the ABP loci (e.g. DFR and ANS) will be necessary.

Further evidence that mutations causing A– phenotypes in *Aquilegia* are likely due to single gene changes comes from crossing studies. Prazmo (1961, 1965), studying the genetics of floral anthocyanin production, using F_2, backcross and F_3 populations, concluded that a single gene was responsible for the A– phenotype in *Aquilegia chrysantha* and *Aquilegia longissima*. We were able to detect only minute amounts of expression of both DFR and ANS in these species throughout floral development, and loss of either enzyme would produce the A– phenotype. Thus, the most parsimonious explanation for Prazmo's (1965) and our results is a single mutation to a common regulatory element of these genes rather than close linkage and independent mutations stopping their expression in flowers. Similarly, Hodges *et al*. (2002) detected a single major quantitative trait locus (QTL) for flower color in a cross between white *Aquilegia pubescens* and red *Aquilegia formosa*. Again, both DFR and ANS have exceptionally low expression in *A. pubescens*. Prazmo (1965) also made extensive studies of segregating populations of white-flowered *A. flabellata* and purple-flowered *Aquilegia ecalcarata*, again concluding that a single gene was responsible for the A– phenotype. Unlike the previous examples, we found that white-flowered *A. flabellata* was down regulated only for CHS suggesting that this lineage has a different cause of the A– phenotype. Significantly, both Prazmo (1965) and Taylor (1984) reported anthocyanin production in F_1 hybrids between white *A. flabellata* and the A– species *A. chrysantha*, *A. longissima* and *A. pubescens* indicating that,

in fact, these lineages do harbour different mutations for the A- phenotype. Future studies mapping the genomic location of the genes and their regulators will be especially useful in substantiating these conclusions.

From this initial survey, it is impossible to determine the critical level for down-regulation to cause the loss of anthocyanin production. Furthermore, this level could be different for each gene. We used an arbitrary cut-off of 50% of maximum expression to categorize genes as down regulated or not. Some A+ taxa have gene expression levels just above this cut-off indicating that this level of expression is sufficient for anthocyanin production. In addition, some species, such as *A. chrysantha*, are classified as down regulated for CHI and F3H yet we detected the biochemical products of these genes at levels found in A+ taxa. Thus, expression in these genes may need to be reduced to very low levels to affect anthocyanin production. It is therefore notable that expression levels of the late genes, DFR, ANS and UF3GT were often completely absent or extremely low making them much more likely to directly affect anthocyanin production. A technical explanation for weak amplification could be nucleotide variation at the primer sites. However, we mostly used degenerate primers, specifically to avoid this problem and when amplification from genomic DNA was successful in A+ taxa it was equally successful in A– taxa suggesting no obvious problems with our PCR primers.

While our data suggest that A– taxa are often convergent due to mutations in trans-regulators, the degree of molecular convergence may not extend beyond this point due to a variety of possible targets. Any of the trans-regulators, bHLH, MYB or WD40, could harbour a mutation that causes loss of anthocyanin production. In addition, a variety of mutations could prevent these factors from binding to their *cis*-counterparts in the ABP genes or prevent the factors from interacting with each other. Lastly, mutations could inhibit the expression patterns of the trans-regulators themselves. Thus, the specific mutations causing A– taxa may not be convergent at a finer level.

One A– lineage (*Aquilegia micrantha*, *Aquilegia barnebyi*) expressed all ABP loci indistinguishable from A+ species indicating that down-regulation has not caused the A– phenotype. In this case, one or more ABP loci may have mutations in their coding sequence, which causes a non-functional enzyme in the pathway. Alternatively, side-branches of the ABP may be directing metabolic flux away from the anthocyanin portion of the pathway. For example, over expression of the side-branch enzyme anthocyanidin reductase (ANR) causes a shift from pink to white flowers in tobacco and from pigmented to nonpigmented seed coats in *Arabidopsis* (Xie *et al*. 2003). The expression of flavones and flavonols at intermediate levels in *A. micrantha* suggests that any structural mutations or competition from side-branch enzymes is located downstream of F3H. Mutations affecting post-transcriptional modifications of ABP loci are yet another possible route to floral anthocyanin losses in these taxa. Given that *A. micrantha* and *A. barnebyi* are sister species, we predict the original loss-of-function mutation to be shared by all populations of both species.

Few other studies have examined the rapid origin of convergent phenotypes at the molecular level. Two notable studies address the loss of body plating in sticklebacks (Colosimo *et al*. 2005) and albinism in cavefish (Protas *et al*. 2006). In the cavefish, Protas *et al*. (2006) show that independent mutations in the pigmentation gene Oca2 are likely responsible for two separate origins of albinism. In contrast, in the sticklebacks, Colosimo *et al*. (2005) examined 15 low-plated populations and found that 14 of these had related alleles of ectodysplasin, Eda, responsible for the low-plated phenotype. Thus, even though neutral markers suggest independent origins of the low-plated phenotype (Rundle *et al*. 2001), it appears that in most cases, a decrease in body plates has evolved through fixation of the same ancestral allele. In plants, hybridization and introgression could lead to patterns resembling those of convergent evolution. For example, introgression of alleles from an A– species into an A+ species could allow the A+ species to evolve into A–. Then the two species would appear to have evolved A– convergently. Although contemporary hybrid zones in *Aquilegia* are rare and geographically restricted, we cannot yet determine if independent regulatory mutations have caused the convergent expression patterns we have found. Future studies of the molecular basis for the A– phenotypes will help distinguish the roles of molecular convergence, parallel fixation of ancestral polymorphisms, and hybridization.

Our comparative expression data are consistent with the importance of regulatory changes during phenotypic evolution, especially during rapid evolutionary change (Britten & Davidson 1969; King & Wilson 1975; Barrier *et al*. 2001), However, for a full understanding of the evolution of this loss-of-function adaptation, a number of additional studies are necessary. In our case, we envision a combination of further expression and comparative sequence analyses. We aim to clone and sequence the regulatory genes to determine if structural or expression differences in them could explain the evolution of most A– species. Similarly, we plan to determine the DNA sequence of the complete ABP structural genes to establish whether these genes are likely to be functional. Identification of the gene sequences from A– species would also allow direct functional assays of the enzymes (Zufall & Rausher 2004). Ultimately, a transformation system will be necessary to establish the causal basis for the evolution of each A– species. The realization of all these approaches will be greatly facilitated by the range of molecular resources that are currently being developed for *Aquilegia*, such as a comprehensive EST database (www.tigr.org/), genetic and physical maps (www.genome.clemson.edu/projects/

aquilegia/), a transformation system and whole-genome sequencing. These tools will not only help us to shed more light on the evolution of flower colour in *Aquilegia* but to address fundamental questions in evolutionary biology, namely what types of mutations are driving ecological speciation and what is the molecular basis of adaptive traits.

Acknowledgements

We thank Elizabeth Zimmer, Douglas Bush, Ji Y. Yang, Kathleen M. Kay, Yael Kisel and five anonymous reviewers for fruitful discussions and helpful comments and suggestions on earlier drafts of the manuscript. Funding was provided by an Olivia Long Converse Fellowship to J.B.W. and National Science Foundation (NSF) (DEB-0309378 and EF-0412727).

Supplementary material

The supplementary material is available from http://www.blackwellpublishing.com/products/journals/suppmat/MEC/MEC3114/MEC3114sm.htm

Table S1 Amino acid alignments for the six structural anthocyanin biosynthetic pathway loci spanning monocots and eudicots and including sequences from *Aquilegia*. Primer lengths are indicated with asterisks at the beginning and the end of the *Aquilegia* sequence. A gap is indicated with a tilde (~).

Table S2 Probability values of correlated changes in expression at DFR and ANS based on independent contrast analysis under four alternative branch-length assignments. Analyses were carried out both with the two horticultural lines *Aquilegia* × *coeruela* 'Origami white' and *Aquilegia flabellata* var. *pumila f. alba* grafted onto the AFLP phylogeny based on the phylogenetic position of their wild relatives (*A. coerulea* and *A. flabellata*, respectively) and also excluding these samples. The significance level ($\alpha = 0.05$) after Bonferroni correction for multiple tests ($N = 15$) is 3.3×10^{-3}.

Fig. S1 (A) Individual, developmental and tissue-specific variation of ABP loci expression in *A. canadensis* (A+) as evaluated by RT-PCR. Five floral stages (GB, green bud, YB, young bud, OB, old bud, PrA, pre-anthesis, PoA, post-anthesis) were analyzed separately in petals (upper lane) and sepals (lower lane) for each of the six ABP loci (see Figure 1 for gene abbreviations) and Actin-1 (ACT). Dotted lines depict the two floral stages (OB, PrA) chosen for the subsequent cross-species comparison. Additionally, ABP loci expression was monitored in anthers of pre-anthesis *A. canadensis* flowers (providing A– tissue from a species that produces A+ petals) of three individual plants. (B) Average expression of actin-standardized anthocyanin biosynthetic pathway loci measured in petals (left panel) and sepals (middle panel) of five floral developmental stages and anthers (right panel) of pre-anthesis flowers of *A. canadensis* (A+) by RT-PCT. Each symbol represents data from a separate individual; lines connect mean values in developmental series and 'x' denotes mean values in anthers. For gene abbreviations see Figure 1.

Fig. S2 (A) Development variation of ABP loci expression in petals of two A+ species [*A. canadensis* (CA) and *A. formosa* (FO)] and one A– species [*A. chrysantha* (CH)] as evaluated by RT-PCR (for gene abbreviations see Figure 1). A+ and A– denote the presence and absence of floral anthocyanins, respectively. Two reactions from the UF3GT locus were unavailable: CA at the green bud stage and FO at the young bud stage. (B) Expression levels of actin-standardized ABP loci across a series of developmental stages in petals of two A+ species [*A. canadensis* (filled diamonds) and *A. formosa* (filled squares)] and one A– species [*A. chrysantha* (open triangles)] based on RT-PCR results (for gene abbreviations see Figure 1). Two reactions from the UF3GT locus were unavailable: CA at the green bud stage and FO at the young bud stage.

Fig. S3 Expression of the core ABP loci across 11 North American *Aquilegia* species and two A– cultivars: *A. laramiensis* (LA), *A. canadensis* (CA), *A. longissima* (LO), *A. pinetorum* (PI), *A. chrysantha* (CH), *A. barnebyi* (BA), *A. micrantha* (MI), *A. coerulea* (CO), *A. pubescens* (PU), *A. flavescens* (FL), *A. formosa* (FO), a white form of the Japanese species *A. flabellata* var. *pumila* f. *alba* (FP), and the white cultivar *A. coerulea* × *hybrida* 'Origami White' (OW). Phylogenetic relationships among the 11 wild species are depicted in the top panel with the dark and light lines indicating A+ and A– species, respectively. All ABP loci and Actin-1 were amplified from old bud-stage flowers (upper row) and pre-anthesis flowers (lower row).

References

Barrier M, Robichaux RH, Purugganan MD (2001) Accelerated regulatory gene evolution in an adaptive radiation. *Proceedings of the National Academy of Sciences, USA*, **98**, 10208–10213.

Beltran M, Jiggins CD, Bull V *et al.* (2002) Phylogenetic discordance at the species boundary: comparative gene genealogies among rapidly radiating *Heliconius* butterflies. *Molecular Biology and Evolution*, **19**, 2176–2190.

Britten RJ, Davidson EH (1969) Gene regulation for higher cells — a theory. *Science*, **165**, 349–357.

Chang S-M, Lu Y, Rausher MD (2005) Neutral evolution of the non-binding region of the anthocyanin regulatory gene Ipmyb1 in *Ipomoea*. *Genetics*, **170**, 1967–1978.

Coberly LC, Rausher MD (2003) Analysis of a chalcone synthase mutant in *Ipomoea purpurea* reveals a novel function for flavonoids: amelioration of heat stress. *Molecular Ecology*, **12**, 1113–1124.

Colosimo PF, Hosemann KE, Balabhadra S *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, **307**, 1928–1933.

Cresko WA, Amores A, Wilson C *et al.* (2004) Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National Academy of Sciences, USA*, **101**, 6050–6055.

Durbin ML, Lundy KE, Morrell PL, Torres-Martinez CL, Clegg MT (2003) Genes that determine flower color: the role of regulatory changes in the evolution of phenotypic adaptations. *Molecular Phylogenetics and Evolution*, **29**, 507–518.

Epperson BK, Clegg MT (1988) Genetics of flower color polymorphism in the common morning glory (*Ipomoea purpurea*). *Journal of Heredity*, **79**, 64–68.

Fehr C, Rausher MD (2004) Effects of variation at the flower-colour A locus on mating system parameters in *Ipomoea purpurea*. *Molecular Ecology*, **13**, 1839–1847.

Felsenstein J (1985) Phylogenies and the comparative method. *The American Naturalist*, **125**, 1–15.

Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thompson JD (2004) Pollination syndromes and floral specialization. *Annual Review of Ecology and Systematics*, **35**, 375–403.

Ffrench-Constant RH, Anthony N, Aronstein K, Rocheleau T, Stilwell G (2000) Cyclodiene insecticide resistance: from molecular to population genetics. *Annual Review of Entomology*, **45**, 449–466.

Garland T, Bennett AF, Rezende EL (2005) Phylogenetic approaches in comparative physiology. *The Journal of Experimental Biology*, **208**, 3015–3035.

Grant V (1952) Isolation and hybridization between *Aquilegia formosa* and *A. pubescens*. *El Aliso*, **2**, 341–360.

Grant V (1992) Foraging ability of rufous hummingbirds on hummingbird flowers and hawkmoth flowers. *Proceedings of the National Academy of Sciences, USA*, **89**, 9400–9404.

Hirschberg J, Cohen M, Harker M *et al.* (1997) Molecular genetics of the carotenoid biosynthesis pathway in plants and algae. *Pure and Applied Chemistry*, **69**, 2151–2158.

Hodges SA, Arnold ML (1994) Columbines: a geographically widespread species flock. *Proceedings of the National Academy of Sciences, USA*, **91**, 5129–5132.

Hodges SA, Whittall JB, Fulton M, Yang JY (2002) Genetics of floral traits influencing reproductive isolation between *Aquilegia formosa* and *Aquilegia pubescens*. *American Naturalist*, **159**, S51–S60.

Hodges SA, Fulton M, Yang JY, Whittall JB (2004) Verne Grant and evolutionary studies of *Aquilegia*. *New Phytologist*, **161**, 113–120.

Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. *The Plant Cell*, **7**, 1071–1083.

King MC, Wilson AC (1975) Evolution at two levels in humans and chimpanzees. *Science*, **188**, 107–116.

Koes RE, Quattrocchio F, Mol JNM (1994) The flavonoid biosynthetic-pathway in plants-function and evolution. *Bioessays*, **16**, 123–132.

Maddison WP, Maddison DR (2005) *MESQUITE: A Modular System for Evolutionary Analysis*, version 1.06 http://mesquiteproject.org.

Majerus MEN, Mundy NI (2003) Mammalian melanism: natural selection in black and white. *Trends in Genetics*, **19**, 585–588.

Mauseth JD (2004) The structure of photosynthetic succulent stems in plants other than cacti. *International Journal of Plant Sciences*, **165**, 1–9.

McCune AR, Carlson RL (2004) Twenty ways to lose your bladder: common natural mutants in zebrafish and widespread convergence of swim bladder loss among teleost fishes. *Evolution and Development*, **6**, 246–259.

Merken HM, Beecher GR (2000) Measurement of food flavonoids by high-performance liquid chromatography: a review. *Journal of Agricultural and Food Chemistry*, **48**, 577–599.

Merken HM, Merken CD, Beecher GR (2001) Kinetics method for the quantitation of anthocyanidins, flavonols, and flavones in foods. *Journal of Agricultural and Food Chemistry*, **49**, 2727–2732.

Midford P, Garland T Jr, Maddison WP (2005) *PDAP: Package of MESQUITE* (version 1.07). http://mesquiteproject.org/pdap-mesquite.

Miller RB (1981) Hawkmoths and the geographic patterns of floral variation in *Aquilegia caerulea*. *Evolution*, **35**, 763–774.

Miller RB (1985) Hawkmoth pollination of *Aquilegia chrysantha* (Ranunculaceae) in southern Arizona. *Southwestern Naturalist*, **30**, 69–76.

Mol J, Grotewold E, Koes R (1998) How genes paint flowers and seeds. *Trends in Plant Science*, **3**, 212–217.

Orr HA (2005) The probability of parallel evolution. *Evolution*, **59**, 216–220.

Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology*, **53**, 673–684.

Prazmo W (1961) Cytogenetic studies on the genus *Aquilegia* II. Crosses between *Aquilegia ecalcarata* Maxim. and *Aquilegia chrysantha* Gray. *Acta Societatis Botanicorum Poloniae*, **30**, 57–77.

Prazmo W (1965) Cytogenetic studies in the genus *Aquilegia*. III. Inheritance of the traits distinguishing different complexes in the genus *Aquilegia*. *Acta Societatis Botanicorum Poloniae*, **34**, 403–437.

Protas ME, Hersey C, Kochanek D *et al.* (2006) Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics*, **38**, 107–111.

Quattrocchio F, Wing J, van der Woude K *et al.* (1999) Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. *Plant Cell*, **11**, 1433–1444.

Rice WR (1988) A new probability model for determining exact *P* values for 2×2 contingency-tables when comparing binomial proportions. *Biometrics*, **44**, 1–14.

Rose TM, Schultz ER, Henikoff JG *et al.* (1998) Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly-related sequences. *Nucleic Acids Research*, **26**, 1628–1635.

Rundle HD, Breden F, Griswold C *et al.* (2001) Hybridization without guilt: gene flow and the biological species concept. *Journal of Evolutionary Biology*, **14**, 868–869.

Simpson GG (1944) *Tempo and Mode in Evolution*. Columbia University Press, New York.

Stebbins GL (1974) *Flowering Plants. Evolution Above the Species Level*. Belknap Press, Cambridge, Massachusetts.

Steyn WJ, Wand SJE, Holcroft DM, Jacobs G (2002) Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. *New Phytologist*, **155**, 349–361.

Strauss SY, Whittall JB (in press) Non-pollinator agents of selection on floral traits. In: *Ecology and Evolution of Flowers* (eds Harder LD, Barrett SCH), Oxford University Press, Oxford.

Taylor RJ (1984) Floral anthocyanins of *Aquilegia* and their relationship to distribution and pollination biology of the species. *Bulletin of the Torrey Botanical Club*, **111**, 462–468.

Whittall JB (2005) *Ecological speciation and convergent evolution in the North American columbine radiation* (*Aquilegia*, Ranunculaceae). PhD, University of California Santa Barbara, California.

Whittall JB, Medina-Marino A, Zimmer EA, Hodges SA (2006) Generating single-copy nuclear gene data in a recent adaptive radiation. *Molecular Phylogenetics and Evolution*, **39**, 124–134.

Whittemore AT (1997) *Aquilegia*. In: *Flora of North America*, **Vol. 3** (ed. Editorial Committee), pp. 249–258. Oxford University Press, New York.

Winkel-Shirley B (2004) It takes a garden. How work on diverse plant species has contributed to an understanding of flavonoid metabolism. *Plant Physiology*, **127**, 1399–1404.

Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA (2003) Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. *Science*, **299**, 396–399.

Zufall RA, Rausher MD (2004) Genetic changes associated with floral adaptation restrict future evolutionary potential. *Nature*, **428**, 847–850.

Justen Whittall is interested in developing and applying comparative phylogenetic approaches to understanding evolutionary patterns of convergence and directionality during plant adaptation and speciation. Claudia Voelckel is interested in combining gene expression and gene divergence analyses to study the molecular basis of adaptations and speciation in plants. Daniel J. Kliebenstein is interested in understanding the evolution of structural diversity in plant secondary metabolites and the resulting pleiotropic effects on plant/biotic interactions as well as human nutrition. The laboratory of Scott A. Hodges studies plant adaptation, speciation and hybridization.
