Transcriptional control of floral anthocyanin pigmentation in monkeyflowers (*Mimulus*)

Yao-Wu Yuan[1,2], Janelle M. Sagawa[1,2], Laura Frost[1], James P. Vela[1] and Harvey D. Bradshaw Jr[1]

[1]Department of Biology, University of Washington, Seattle, WA 98195, USA; [2]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, CT 06269, USA

Authors for correspondence:
Yao-Wu Yuan
Tel: +1 860 486 3469
Email: yuan.colreeze@gmail.com

Harvey D. Bradshaw Jr
Tel: +1 206 616 1796
Email: toby@uw.edu

Summary

• A molecular description of the control of floral pigmentation in a multi-species group displaying various flower color patterns is of great interest for understanding the molecular bases of phenotypic diversification and pollinator-mediated speciation.

• Through transcriptome profiling, mutant analyses and transgenic experiments, we aim to establish a 'baseline' floral anthocyanin regulation model in *Mimulus lewisii* and to examine the different ways of tinkering with this model in generating the diversity of floral anthocyanin patterns in other *Mimulus* species.

• We find one *WD40* and one *bHLH* gene controlling anthocyanin pigmentation in the entire corolla of *M. lewisii* and two *R2R3-MYB* genes, *PELAN* and *NEGAN*, controlling anthocyanin production in the petal lobe and nectar guide, respectively. The autoregulation of *NEGAN* might be a critical property to generate anthocyanin spots. Independent losses of *PELAN* expression (via different mechanisms) explain two natural yellow-flowered populations of *M. cardinalis* (typically red-flowered). The *NEGAN* ortholog is the only anthocyanin-activating *MYB* expressed in the *M. guttatus* flowers.

• The mutant lines and transgenic tools available for *M. lewisii* will enable gene-by-gene replacement experiments to dissect the genetic and developmental bases of more complex floral color patterns, and to test hypotheses on phenotypic evolution in general.

Introduction

Flower color produces some of the most beautiful displays in nature, and serves an important function in the ecology and evolution of plants by attracting animal pollinators (Glover, 2007; Davies *et al.*, 2012). Many closely related plant species display dramatically different floral color patterns, and in some cases the flower color or pattern change may have produced a pollinator shift that leads to reproductive isolation and speciation (Bradshaw & Schemske, 2003; Hoballah *et al.*, 2007; Streisfeld *et al.*, 2013). A molecular description of the control of floral pigmentation in a multi-species group displaying various flower colors and patterns would, therefore, be of great interest for understanding the molecular bases underlying phenotypic diversification, plant–pollinator interactions and pollinator-mediated speciation. The monkeyflower genus, *Mimulus*, provides an excellent study system for such an endeavor.

The genus *Mimulus* contains 160–200 species that exhibit astonishing flower color variation (Fig. 1) and has been the subject of intensive ecological and evolutionary studies for over 60 yr (Hiesey *et al.*, 1971; Beardsley *et al.*, 2004; Wu *et al.*, 2008). In the past decade a wealth of genomic and genetic resources as well as functional tools have been developed for multiple species in the genus (Wu *et al.*, 2008; Cooley *et al.*, 2011; Hellsten *et al.*, 2013; Streisfeld *et al.*, 2013; Yuan *et al.*, 2013a,b), enabling

in-depth genetic and developmental analyses of this traditionally ecological and evolutionary model system.

Two major pigment types determine flower color in *Mimulus*: anthocyanins are responsible for the pink/purple color and carotenoids for the yellow (Vickery & Olson, 1956; Hiesey *et al.*, 1971; Streisfeld & Kohn, 2005; Cooley & Willis, 2009; Yuan *et al.*, 2013b). A combination of the two pigment types often result in red color, as in the petal lobes of *M. cardinalis* and the spots on the corolla throat of *M. guttatus* (Fig. 1).

The core enzymes involved in both anthocyanin and carotenoid biosyntheses have been well characterized (reviewed in Grotewold, 2006; Glover, 2007; Ruiz-Sola & Rodriguez-Concepcion, 2012). The diversity of flower color in nature is largely determined by when and where these enzymes are expressed (i.e. transcriptional regulation of these enzymes) (Schwinn *et al.*, 2006; Glover, 2007; Shang *et al.*, 2011; Martins *et al.*, 2013). While the transcriptional control of anthocyanin biosynthetic enzymes has been elucidated in multiple plant species (Paz-Ares *et al.*, 1987; Ludwig *et al.*, 1989; Goodrich *et al.*, 1992; Quattrocchio *et al.*, 1998; Walker *et al.*, 1999; Borevitz *et al.*, 2000; Spelt *et al.*, 2000; Schwinn *et al.*, 2006; Albert *et al.*, 2011, 2014; reviewed in Koes *et al.*, 2005; Hichri *et al.*, 2011; Davies *et al.*, 2012), the regulation of carotenoid biosynthetic pathway in flowers is little known (Grotewold, 2006; Glover, 2007; Ruiz-Sola & Rodriguez-Concepcion, 2012). As such, anthocyanin pigmentation provides
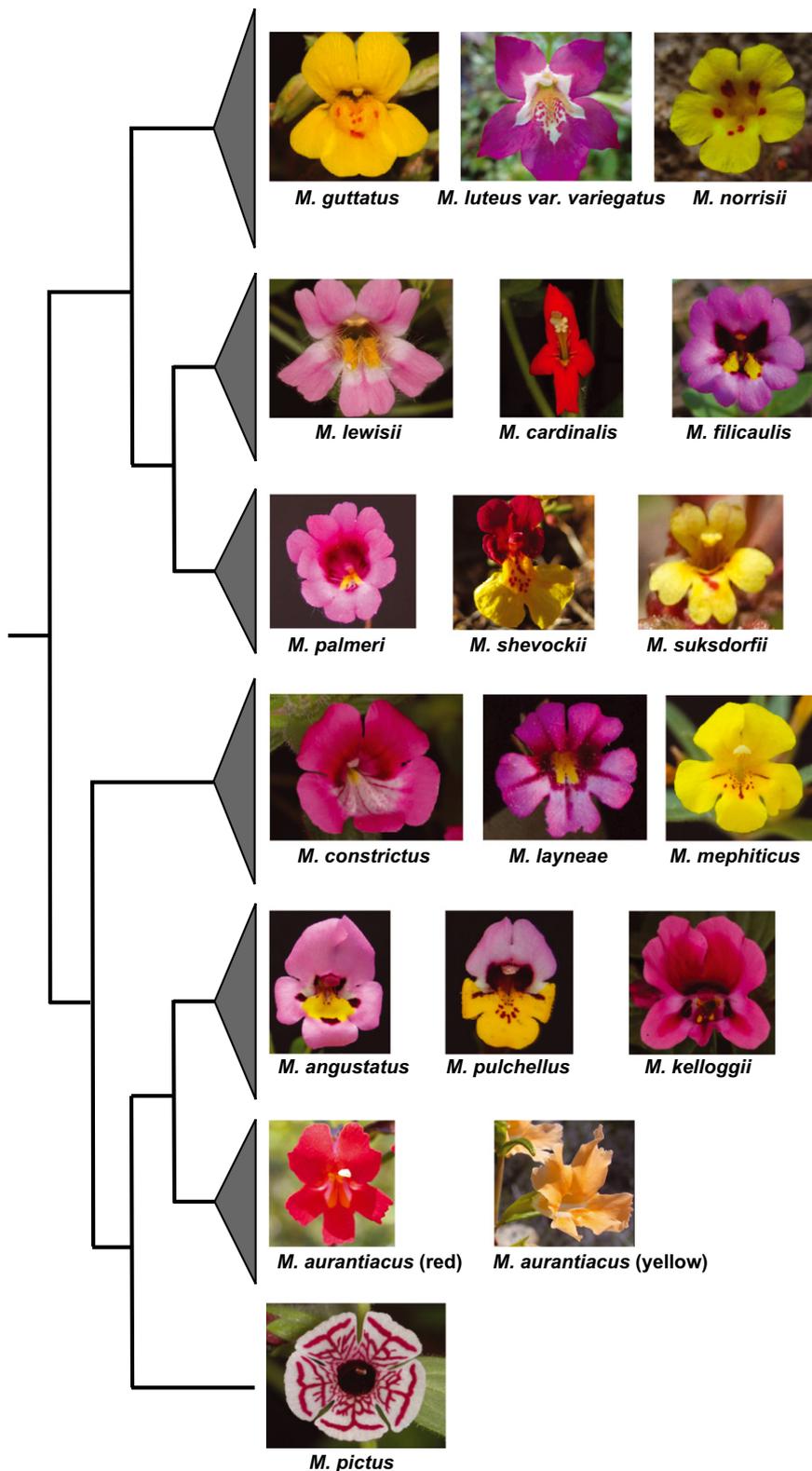
Fig. 1 Natural flower color variation among *Mimulus* species. Shown on the left is a schematic illustration of the phylogenetic relationships among major *Mimulus* clades, based on Beardsley *et al.* (2004) and Grossenbacher & Whittall (2011); on the right are representative species of each clade. Images of *M. norrisii*, *M. filicaulis*, *M. palmeri*, *M. shevockii*, *M. suksdorfii*, *M. layneae*, *M. mephiticus*, *M. angustatus* and *M. pulchellus* were provided by Dena Grossenbacher; images of *M. aurantiacus* were provided by Matt Streisfeld; the remaining images were taken by Y-W.Y.

a more suitable platform at the moment to understand the genetic and developmental bases of natural variation between different species (Sobel & Streisfeld, 2013).

The anthocyanin biosynthetic pathway (ABP) contains at least six essential structural genes encoding core enzymes:

Chalcone synthase (*CHS*), *Chalcone isomerase* (*CHI*), *Flavonoid 3-hydroxylase* (*F3H*), *Dihydroflavonol 4-reductase* (*DFR*), *Anthocyanidin synthase* (*ANS*) and *UDP-3-O-glucosyltransferases* (*UF3GT*). In maize (*Zea mays*), *Petunia* and *Arabidopsis*, a highly conserved MYB-bHLH-WD40 (MBW) protein complex

has been shown to coordinately activate all or some of the ABP structural genes (Paz-Ares *et al.*, 1987; Ludwig *et al.*, 1989; de Vetten *et al.*, 1997; Quattrocchio *et al.*, 1999; Walker *et al.*, 1999; Borevitz *et al.*, 2000; Spelt *et al.*, 2000; Zhang *et al.*, 2003; Carey *et al.*, 2004). In *Antirrhinum majus* flowers, three closely related R2R3-MYBs (Rosea1, Rosea2 and Venosa) and a bHLH (Delila) protein coordinately activate at least four ABP genes (*F3H*, *DFR*, *ANS* and *UF3GT*) (Martin *et al.*, 1991; Goodrich *et al.*, 1992; Schwinn *et al.*, 2006), but the WD40 component has yet to be identified. The MYB and bHLH proteins represent the two largest transcription factor families in flowering plants (Feller *et al.*, 2011). The anthocyanin-activating MYBs and bHLHs belong to subgroup 6 of the R2R3-MYB family and subgroup IIIf of the bHLH family, respectively (Stracke *et al.*, 2001; Heim *et al.*, 2003; Feller *et al.*, 2011). More recently, a group of single-repeat R3-MYBs has been shown to negatively regulate ABP gene expression by competing with the R2R3-MYB activators for the limited supply of bHLH proteins (Zhu *et al.*, 2009; Nakatsuka *et al.*, 2013; Yuan *et al.*, 2013a; Albert *et al.*, 2014).

The conserved nature of the ABP enzymes and their MBW regulators across flowering plants has enabled a number of investigations on the genetic control of floral anthocaynin pigmentation in nonmodel organisms, including *Phalaenopsis* orchids (Ma *et al.*, 2009), Asiatic hybrid lilies (*Lilium* spp.) (Yamagishi *et al.*, 2010, 2014), morning glories (*Ipomoea* spp.) (Des Marais & Rausher, 2010), *Iochroma* spp. (Smith & Rausher, 2011), *Phlox drummondii* (Hopkins & Rausher, 2011) and *Clarkia gracilis* (Martins *et al.*, 2013). These studies provide valuable information on the potential players involved in flower color diversification in a wide range of angiosperm lineages. However, due to the lack of genetic resources or/and functional tools, reaching a deeper understanding of the precise molecular bases and developmental mechanisms that generate floral anthocyanin patterns (e.g. spots, stripes) or cause flower color variation between species, remains a formidable task in most of these systems.

Here we describe the major transcriptional regulators of ABP genes in the pink flowered *Mimulus lewisii* (Fig. 1), an emerging model system particularly suitable for studying the developmental genetics of ecologically important floral traits (Yuan *et al.*, 2013a,b). Through transcriptome profiling, mutant analyses and transgenic experiments, we find one *WD40* (*MlWD40a*) and one *bHLH* (*MlANbHLH1*) gene controlling anthocyanin biosynthesis in the entire corolla of *M. lewisii*; one *R2R3-MYB*, named *Petal Lobe Anthocyanin* (*PELAN*), controlling anthocyanin biosynthesis in the petal lobe, and another *R2R3-MYB*, named *Nectar Guide Anthocyanin* (*NEGAN*), controlling anthocyanin spot formation in the nectar guide. *NEGAN*, but not *PELAN*, is involved in an autoregulatory feedback loop, which might be a critical property required for the 'spot' pattern formation. Furthermore, using the *M. lewisii* model we demonstrate that two independent losses of *PELAN* expression (via different mechanisms) explain the yellow flower color of two natural populations of *M. cardinalis* (usually red-flowered), which is the sister species of *M. lewisii*. Our *M. lewisii* model also explains the *M. guttatus* anthocyanin pattern by successfully predicting the *NEGAN*

ortholog as the only anthocyanin-activating *MYB* expressed in the predominantly yellow *M. guttatus* flowers.

Materials and Methods

Plant materials and growth conditions

The *Mimulus lewisii* inbred line LF10 and *M. cardinalis* inbred line CE10 were described in Yuan *et al.* (2013a). The *M. lewisii boo* mutants were generated by ethyl methanesulfonate (EMS) mutagenesis in the LF10 background (Owen & Bradshaw, 2011). Seeds of two natural yellow-flowered *M. cardinalis* strains, SM and CI, collected from the Siskiyou Mountains of Oregon (US) and Cedros Island (Baja California, Mexico), respectively, were provided by Bob Vickery (University of Utah). Seeds of *M. guttatus* inbred line IM767 were provided by John Willis (Duke University). Plants were grown in the University of Washington and University of Connecticut glasshouses under similar conditions as described in Yuan *et al.* (2013a).

Transcriptome sequencing and analyses

In order to obtain a comprehensive view of the expression profile of the ABP genes and their putative transcriptional regulators in the *M. lewisii* LF10 flowers, we isolated total RNA from the corolla of 15-mm flower buds (3 d before opening) for transcriptome sequencing. The 15-mm corolla stage is the intermediate stage between when anthocyanins first become visible (10-mm, 6 d before opening) and flower opening (Yuan *et al.*, 2013a) – not too early to capture ABP gene expression and not too late to capture the expression of their transcriptional regulators.

The RNA-Seq library was prepared at the University of North Carolina High-Throughput Sequencing Facility (UNC-HTSF) using the Illumina (San Diego, CA, USA) TruSeq RNA Sample Preparation Kit v2. Briefly, mRNA was first purified from 2 μg of total RNA using oligo-dT attached magnetic beads, and then cleaved into 200-bp pieces under elevated temperature. The resulting RNA fragments were primed with random hexamers and were reverse transcribed into first strand cDNA, followed by second strand cDNA synthesis. The double strand cDNA was end repaired and A-tailed, and then ligated to adapters for PCR enrichment to generate the final cDNA library for Illumina sequencing at the UNC-HTSF.

The resulting *c.* 85 million 100-bp paired-end RNA-Seq reads (NCBI Sequence Read Archive PRJNA232780: SRX403785) were assembled into 80 602 contigs (N50 = 1.4 kb; average length = 834 bp) using CLC Genomics Workbench (Qiagen, Venlo, Netherlands) with default *de novo* assembly parameters. We then mapped the RNA-Seq reads to the *de novo* transcriptome assembly (available from http://www.eeb.uconn.edu/people/yuan/resources) to determine the RPKM expression value (Mortazavi *et al.*, 2008) of each transcript using the CLC Genomics Workbench 'RNA-Seq Analysis' tool, with the minimum read length fraction set to 0.9 and minimum similarity set to 0.97. Previously characterized ABP genes and their MBW regulators from *Arabidopsis* were used as queries (Table 1) to retrieve the

Table 1 Gene expression profile from the *Mimulus lewisii* LF10 15-mm corolla transcriptome

Query	Gene	GenBank acc. no.	Expression value (RPKM)
TT4 (AT5G13930)	*MlCHSa*	KJ011133	2696.65
	MlCHSb	KJ595581	0
	MlCHSc	KJ595582	0
TT5 (AT3G55120)	*MlCHI*	KJ011134	465.19
TT6 (AT3G51240)	*MlF3Ha*	KJ011135	368.86
	MlF3Hb	KJ595583	49.33
TT3 (AT5G42800)	*MlDFR*	KJ011136	188.31
TT18 (AT4G22880)	*MlANS*	KJ011137	305.36
UF3GT (AT5G54060)	*MlUF3GT*	KJ011138	571.61
TTG1 (AT5G24520)	*MlWD40a*	KJ011139	30.34
	MlWD40b	KJ011140	3.08
	MlWD40c	KJ011141	5.78
GL3 (AT5G41315),	*MlANbHLH1*	KJ011142	75.95
EGL3 (AT1G63650),	*MlANbHLH2*	KJ011143	12.55
TT8 (AT5G24520)	*MlANbHLH3*	KJ789366	0
PAP1 (AT1G56650),	*PELAN*	KJ011144	124.83
PAP2 (AT1G66390)	*NEGAN*	KJ011145	8.56
	PELAN-L1	KJ595584	0
	PELAN-L2	KJ595585	0
	PELAN-L3	KJ595586	0

The genes with '0' RPKM values are found in the genome assembly, but not in the transcriptome.

corresponding transcripts from the LF10 *de novo* transcriptome assembly by TBLASTN searches (Altschul *et al.*, 1997). To complement the transcriptome data, we also used the same query sequences to search against the genome assembly generated in a previous study (Yuan *et al.*, 2013a). The genomic copies of these genes were annotated and the sequences were deposited in GenBank (accession numbers are listed in Table 1).

Phylogenetic analyses

Multiple sequence alignments of R2R3-MYB and bHLH proteins were performed using MUSCLE (Edgar, 2004). Conserved regions of the alignments were selected using Gblocks (Talavera & Castresana, 2007) for subsequent phylogenetic analyses. Maximum likelihood (ML) analyses were conducted using RAxML 7.0.4 (Stamatakis, 2006), with the JTT amino acid substitution matrix and the GAMMA model of rate heterogeneity. Clade support was estimated by 200 bootstrap replicates.

Expression analyses by RT-PCR

Total RNA was isolated using the Spectrum Plant Total RNA Kit (Sigma-Aldrich) and then treated with amplification grade DNaseI (Invitrogen). cDNA was synthesized from 1 µg of the DNase-treated RNA using the SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen), then diluted 40-fold before PCR. The *Mimulus* ortholog of *At5g25760/Ubiquitin-Conjugating Enzyme (UBC)* was used as a reference gene as described in Yuan *et al.* (2013a). Gene-specific primers used for RT-PCR are listed in Supporting Information Table S1. Quantitative RT-PCR was performed using iQ SYBR Green Supermix

(Bio-Rad) in a CFX96 Touch Real-Time PCR Detection System (Bio-Rad). Samples were amplified for 40 cycles of 95°C for 15 s and 60°C for 30 s. Reactions were run with three biological replicates and two technical replicates. Amplification efficiencies for each primer pair were determined using critical threshold values obtained from a dilution series (1 : 4, 1 : 20, 1 : 100, 1 : 500).

Candidate gene sequencing of the *boo* mutants

The transcriptome analyses revealed one *WD40*, one *bHLH* and one *R2R3-MYB* as potential candidate genes encoding the major transcriptional regulators. To examine whether these genes harbor mutations in the four *boo* mutant lines that were identified as potential loss-of-function mutants of the MBW complex (see the Results section), full-length coding DNA sequences (CDS) of these candidate genes were amplified from the *boo* cDNAs. PCR products were treated with ExoSAP-IT (USB/Affymetrix, Santa Clara, CA, USA) and sequenced using the BigDyeTerminator v3.1 system (Applied Biosystems, Foster City, CA, USA) following the manufacturers' protocols. Primers used for amplification and sequencing are listed in Table S2.

Transgenic experiments

RNAi plasmids were constructed with a 133–350-bp fragment amplified from the coding regions of *MlWD40a*, *MlANbHLH1*, *PELAN* and *NEGAN* (Table S3), essentially following the protocol described in Yuan *et al.* (2013a). To ensure target specificity, the fragment included in each RNAi plasmid was BLASTed against the LF10 genome assembly with an *E*-value cutoff of 0.1 so that no other genomic regions perfectly match this fragment for a contiguous block longer than 16 bp. For *PELAN*, it was not possible to find a contiguous region longer than 100 bp in the coding region fulfilling this criterion. Therefore, we connected two shorter fragments (75 and 68 bp) with the required specificity by bridge PCR (Table S3).

In order to test whether *NEGAN* is self-activated, we generated an over-expression plasmid by cloning the 804-bp full-length *NEGAN* CDS (without the stop codon; Table S3) into the pEarleyGate 103 vector (Earley *et al.*, 2006; Arabidopsis Biological Resource Center, CD3-685), following Earley *et al.* (2006). This vector drives the expression of the transgene by the CaMV 35S promoter.

The final plasmid constructs were verified by sequencing and then transformed into *Agrobacterium tumefaciens* strain GV3101 for subsequent plant transformation, as described in Yuan *et al.* (2013a).

Results

Identification of the putative ABP structural genes and MBW regulatory genes by BLAST searches

BLAST searches against the LF10 15-mm corolla transcriptome identified one copy of each of the six ABP structural genes (*MlCHSa*, *MlCHI*, *MlF3Ha*, *MlDFR*, *MlANS* and *MlUF3GT*)

with high RPKM expression values, and an additional copy of the *F3H* gene (*MlF3Hb*) with relatively low RPKM value (Table 1). Searches against the LF10 genome assembly confirmed the single copy of *CHI*, *DFR*, *ANS* and *UF3GT*, and the two paralogs of *F3H*, but revealed two additional paralogs of *CHS* (*MlCHSb* and *MlCHSc*). RT-PCR across four corolla developmental stages (5-, 10-, 15- and 20-mm) suggests that *MlCHSb* and *MlCHSc* are not expressed in the corolla (Supporting Information Fig. S1a), explaining the absence of these two copies in the transcriptome. Furthermore, the RT-PCR experiments showed that *MlF3Ha* has much higher expression levels than *MlF3Hb* in the 10- and 15-mm stages. *MlF3Hb* expression was not detectable until later stages (15- and 20-mm) (Fig. S1b). Because anthocyanin pigments already become visible in the 10-mm corolla (Yuan *et al.*, 2013a), we reasoned that *MlF3Ha*, not *MlF3Hb*, plays a primary role in anthocyanin biosynthesis in the corolla. Taking these results together, we concluded that *MlCHSa*, *MlCHI*, *MlF3Ha*, *MlDFR*, *MlANS* and *MlUF3GT* are

the six key ABP structural genes responsible for LF10 corolla anthocyanin pigmentation.

Similarly, a combination of transcriptome and genome BLAST searches retrieved five paralogs of subgroup 6 *R2R3-MYB*s, three subgroup IIIf *bHLH*s paralogs and three *WD40* paralogs. Although flowering plant genomes typically harbor hundreds of *R2R3-MYB* and *bHLH* genes (Stracke *et al.*, 2001; Heim *et al.*, 2003; Pires & Dolan, 2010; Feller *et al.*, 2011), identification of the anthocyanin-activating subgroup 6 R2R3-MYBs and subgroup IIIf bHLHs are facilitated by the 'signature' amino acid sequence motif, '[R/K]P[R/Q]PRx[F/L]' (Stracke *et al.*, 2001; Fig. S2) and 'NGxIKTRKxxQxxExxx[D/E]xxxLxRSxQLREL YESLxxxE' (Pires & Dolan, 2010; Fig. S3), that clearly distinguishes these two subgroups, respectively. Phylogenetic analyses including the *M. lewisii* sequences and previously characterized R2R3-MYB and bHLH sequences corroborated our identification based on the signature motifs (Fig. 2). The WD40 protein sequences are extremely different from other WD-repeat
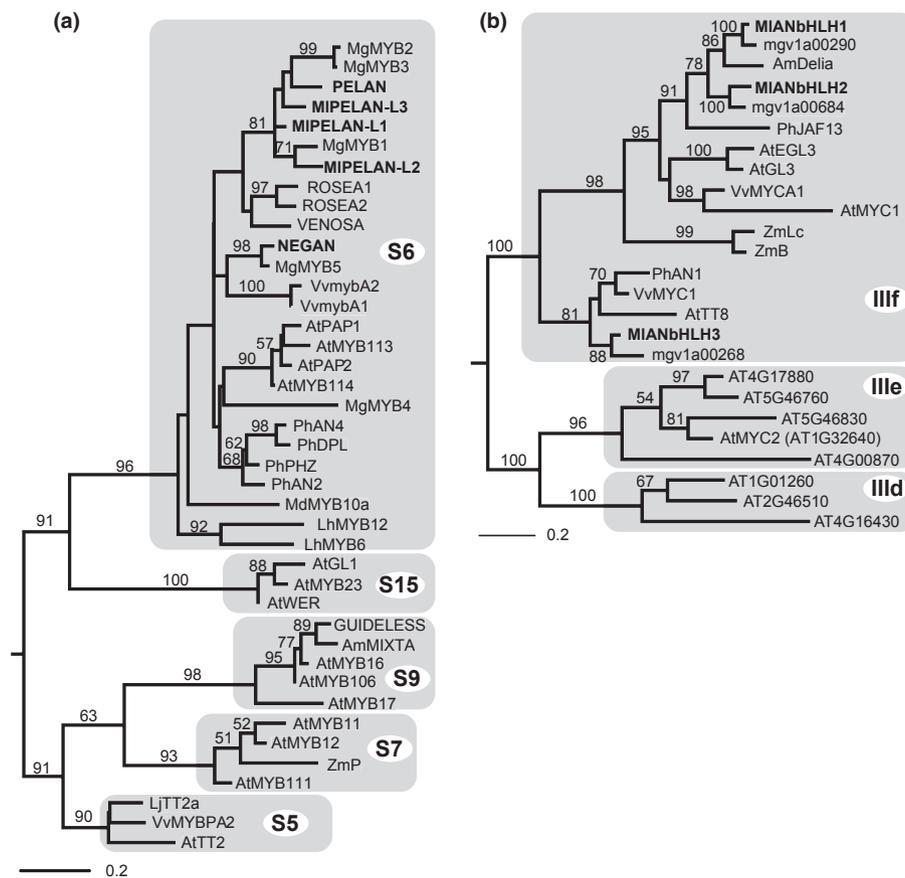
Fig. 2 Maximum likelihood (ML) phylogenies of R2R3-MYB (subgroup 6 and related subgroups) proteins (a) and bHLH (subgroup IIIf and related subgroups) proteins (b). Subgroup classifications followed Stracke *et al.* (2001) and Heim *et al.* (2003). Trees are rooted by midpoint rooting. Bootstrap support values > 50% are indicated along the branches. *Mimulus lewisii* sequences characterized in this study are highlighted in bold and have been deposited in GenBank (accession numbers see Table 1). All *M. guttatus* sequences were retrieved from Phytozome (http://www.phytozome.net), including MgMYB1-5 (MgMYB1: mgv1a023671m; MgMYB2: mgv1a024703m; MgMYB3: mgv1a024996m; MgMYB4: mgv1a025765m; MgMYB5: mgv1a019326m) and the three bHLHs (mgv1a00290, mgv1a00684, mgv1a00268). All *Arabidopsis* sequences were retrieved from the TAIR site (http://www.arabidopsis.org/); Other sequences were retrieved from GenBank as follows: *Antirrhinum majus* ROSEA1 (DQ275529); ROSEA2 (DQ275530); Venosa (DQ275531); AmMIXTA (X79108); AmDelia (AAA32663); *Lilium hybrid* LhMYB6 (AB534587); LhMYB12 (AB534586); *Lotus japonicus* TT2a (BAG12893); *Malus × domestica* MdMYB10a (DQ267897); *M. lewisii* GUIDELESS (KC139356); *Petunia × hybrida* PhAN2 (AF146702); PhAN4 (HQ428105); PhDPL (HQ116169); PhPHZ (HQ116170); PhAN1 (AAG25928); PhJAF13 (AAC39455); *Vitis vinifera* VvmybA1 (BAD18977); VvmybA2 (BAD18978); VvMYBPA2 (ACK56131); VvMYC1 (EU447172); VvMYCA1 (EF193002); *Zea mays* ZmP (P27898); ZmB (CAA40544); ZmLc (AAA33504).

containing proteins, but highly conserved among WD40 members (Fig. S4). As such, the identification of WD40 sequences is straightforward. The transcriptome profiling suggests that only two of the five *R2R3-MYB* paralogs are expressed in the LF10 corolla; one has much higher expression value than the other. Likewise, only one *bHLH* paralog and one *WD40* paralog are predominantly expressed in the corolla (Table 1). RT-PCR across different corolla developmental stages confirmed the transcriptome results (Fig. S1c–e).

boo1, *boo3*, *boo5* and *boo14* are putative loss-of-function mutants of the MBW regulatory complex

In order to determine whether these putative MBW genes retrieved from BLAST searches are actual transcriptional regulators of anthocyanin pigmentation in the *M. lewisii* flower, we set out to identify loss-of-function mutants of the MBW regulatory complex. We screened 12 recessive *boo* mutants with white flowers for coordinated downregulation of ABP structural genes. The rationale is that loss-of-function of one structural gene in the pathway should not affect the expression of other ABP genes, whereas loss-of-function of the MBW regulatory complex will lead to downregulation of multiple ABP genes simultaneously. Subsequent RT-PCR of the six key structural genes showed dramatic downregulation of *MlF3Ha*, *MlDFR* and *MlANS* in four of the 12 *boo* mutants (Fig. 3). Pair-wise complementation crosses suggest that these four mutant lines belong to three complementation groups: *boo3*, *boo1/boo5* and *boo14*.

Among the three complementation groups, *boo3* has the most specific phenotypic effect (i.e. least pleiotropy), only lacking the anthocyanin pigments on the petal lobe. The anthocyanin spots in the nectar guide and the purple color of the stem base remain unaffected. The seed coat color of *boo3* is also indistinguishable from the wild-type LF10 (Fig. 4), being deep brown, presumably due to proanthocyanidin accumulation (Lepiniec *et al.*, 2006). The *boo1/boo5* lines show the most pleiotropy: petal lobes are

white, anthocyanin spots in the nectar guide are completely absent, the stem base has no purple color and seed coat is pale yellow (Fig. 4). *boo14* shows intermediate phenotypes, with no anthocyanins in the petal lobes or stem base, weak anthocyanin spots towards the base of the nectar guide, but having a seed coat color indistinguishable from the wild-type (Fig. 4).

boo3, *boo1/boo5* and *boo14* correspond to the *R2R3-MYB*, *WD40* and *bHLH* genes, respectively

The different degree of pleiotropy displayed by the *boo3*, *boo1/boo5* and *boo14* mutants suggest that they may correspond to the three different classes of transcriptional regulators: *WD40* mutations are usually the most pleiotropic and *R2R3-MYB* mutations are the least pleiotropic (Zhang *et al.*, 2003; Koes *et al.*, 2005). The *R2R3-MYB*, *bHLH* and *WD40* genes with the highest expression level in the corolla were considered as the most promising candidate gene underlying *boo3*, *boo14* and *boo1/boo5*, and were named *Petal Lobe Anthocyanin* (*PELAN*), *MlANbHLH1* and *MlWD40a*, respectively (Table 1). To test this idea, we sequenced the full-length CDS of the three candidate genes for all four *boo* mutants. In each *boo* mutant line only one of the three candidate genes harbored a mutation. The *boo3* mutant has a premature stop codon at the end of the first exon of *PELAN* (Figs 5a,S2); *boo1* and *boo5* have an amino acid replacement in a highly conserved site and a premature stop codon, respectively, in *MlWD40a* (Figs 5b, S4), consistent with the complementation test indicating that they are allelic; *boo14* has a mutation in an intron/exon junction leading to nonsplicing of the last intron of *MlANbHLH1* (Figs 5c, S3). These results support our hypothesis that *boo3*, *boo1/boo5* and *boo14* correspond to the *R2R3-MYB* gene *PELAN*, the *WD40* gene *MlWD40a* and the *bHLH* gene *MlANbHLH1*, respectively.

A second *R2R3-MYB* controls anthocyanin spot formation in the nectar guide

The fact that both *MlWD40a* and *MlANbHLH1* mutations affect the anthocyanin spots in the nectar guide, whereas the *PELAN* mutation does not, suggests that there must be another *R2R3-MYB* responsible for anthocyanin spot formation in the nectar guide. The transcriptome and RT-PCR experiments did show one other *R2R3-MYB* paralog expressed in the corolla, although the RPKM expression value of this gene is much lower than *PELAN* (8.56 vs 124.83; Table 1). However, if this gene is only expressed in the nectar guide, the relatively low expression value estimated from the entire corolla can be explained by the fact that the nectar guide accounts for only a small proportion (*c.* 10%) of the corolla tissue. To test this hypothesis, we dissected the 15-mm corolla into petal lobes and nectar guide and isolated RNA from each tissue type. RT-PCR showed that *PELAN* is predominantly expressed in the petal lobe (Fig. 5d), whereas the other *R2R3-MYB* paralog is expressed exclusively in the nectar guide. These results strongly suggest that this second *R2R3-MYB*, designated as *Nectar Guide Anthocyanin* (*NEGAN*), controls the anthocyanin spot formation in the nectar guide.

Fig. 3 Semi-quantitative RT-PCR of the structural anthocyanin biosynthetic genes in the *Mimulus lewisii boo* mutants. *MlF3Ha*, *MlDFR* and *MlANS* are dramatically downregulated in these mutants; *MlUF3GT* is also downregulated, but to a lesser extent; *MlCHSa* and *MlCHI* are not affected. *MlUBC* is shown as a reference gene. PCR cycle numbers are shown after the gene names. WT, wild-type.
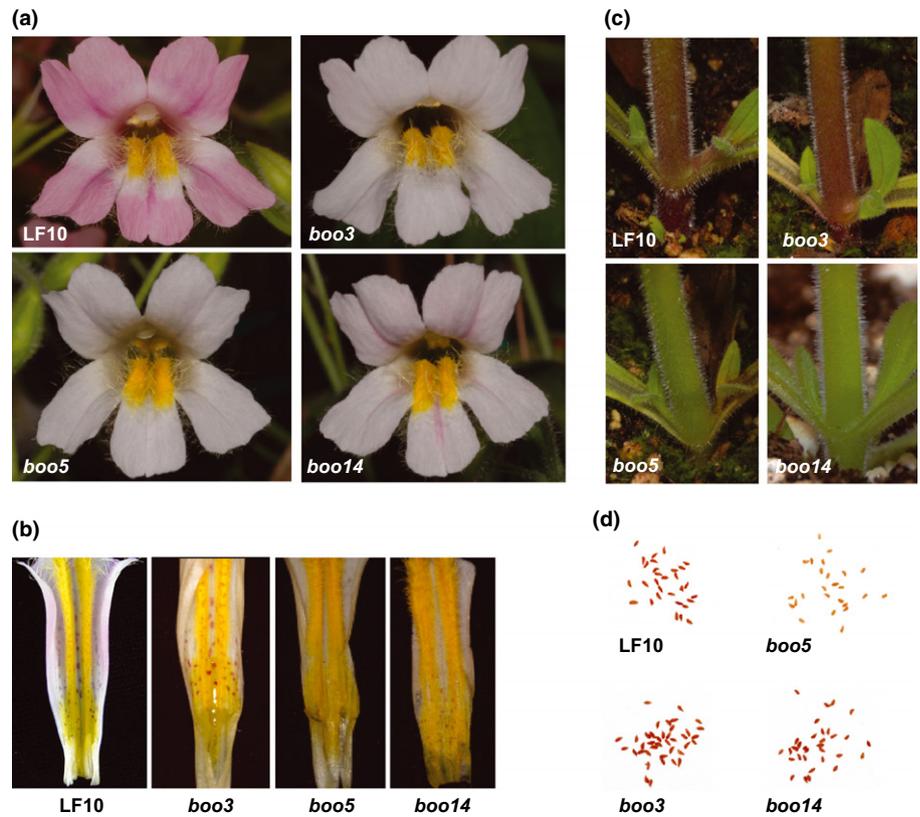
Fig. 4 Phenotypes of the wild-type *Mimulus lewisii* LF10 and the *boo3*, *boo5* and *boo14* mutants. (a) Petal lobe color. (b) Anthocyanin spots in the nectar guide. (c) Color of the stem base. (d) Color of the seed coat.
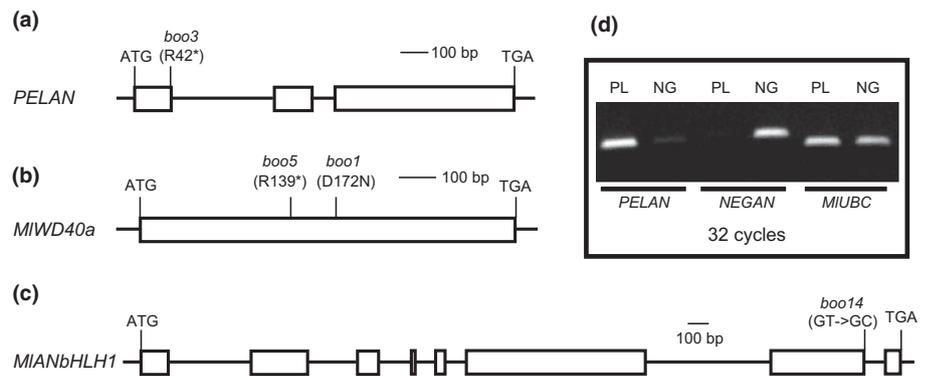
Fig. 5 Mutation characterization of the *Mimulus lewisii boo* mutants. (a) *boo3* has a premature stop codon in the *PELAN R2R3-MYB* gene. (b) *boo1* and *boo5* have a nonsynonymous mutation (D172N) and a premature stop codon, respectively, in the *MlWD40a* gene. (c) *boo14* has a mutation in an intron/exon junction in *MlANbHLH1*, leading to nonsplicing of the last intron. (d) Expression pattern of the *PELAN MYB* and the *NEGAN MYB* in the wild-type petal lobe (PL) and nectar guide (NG).

RNAi transgenic lines accurately reproduce *boo* phenotypes

In order to further verify the function of these MBW regulators, we built gene-specific RNAi constructs to knock down the expression of each of the four players identified above, in the wild-type LF10 background. We generated 4, 92, 12 and 14 independent RNAi transgenic lines for *PELAN*, *NEGAN*, *MlANbHLH1* and *MlWD40a*, respectively. For all genes but *MlANbHLH1*, more than half of the transgenic lines display a strong phenotype indistinguishable from the corresponding *boo* mutants; the remaining lines usually show an intermediate phenotype (Fig. 6). Even for *MlANbHLH1*, five of 12 independent lines have strong phenotypes indistinguishable from *boo14*.

The strong *PELAN* RNAi lines have white petal lobes, but have normal nectar guide anthocyanin spots (Fig. 6a), purple stem base and deep brown seed coat color, accurately

phenocopying *boo3*. The strong *NEGAN* RNAi lines have no anthocyanin spots in the nectar guide, but are indistinguishable from LF10 in all other traits, including petal lobe color (Fig. 6b). The strong *MlANbHLH1* lines have white petal lobes, weak anthocyanin spots towards the base of the nectar guide (Fig. 6c), a green stem base, but normal seed coat color, just like the *boo14* mutant. The fact that both *boo14* and the *MlANbHLH1* RNAi lines have weak anthocyanin spots in the bottom part of the nectar guide suggests that there probably exists another functionally redundant *bHLH* paralog (e.g. *MlANbHLH2*, Table 1) with low expression level in the nectar guide. *MlWD40a* knock-down affects all of the aforementioned traits, as in *boo5*. Together with the mutant analyses, these transgenic results support the model that *MlANbHLH1* and *MlWD40a* regulate anthocyanin biosynthesis in the entire corolla, while *PELAN* and *NEGAN* act more specifically in different parts of the corolla (Fig. 7).
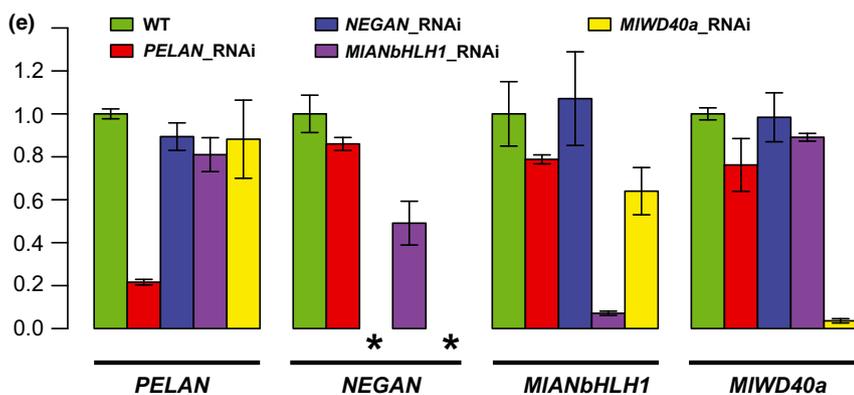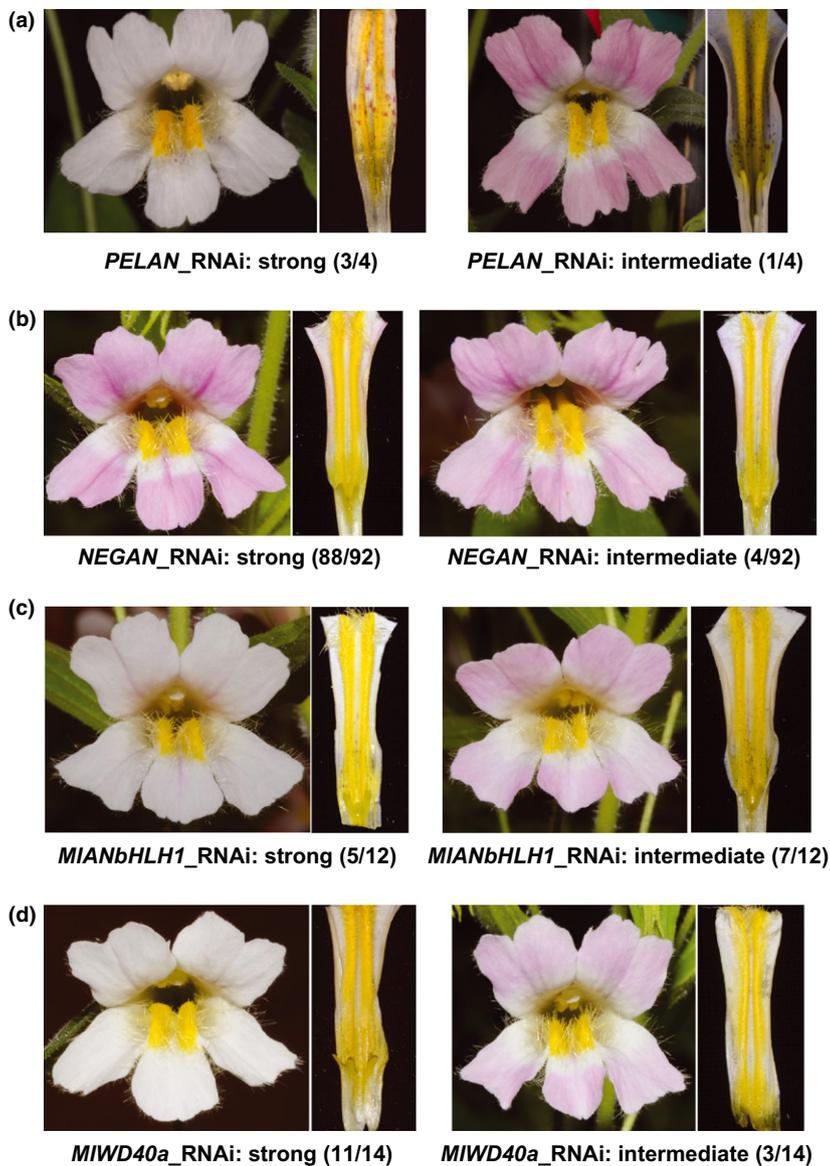
Fig. 6 Phenotypic and molecular characterization of the RNAi transgenic lines of *PELAN* (a), *NEGAN* (b), *MlANbHLH1* (c) and *MlWD40a* (d) in the *Mimulus lewisii* LF10 background. The proportion of transgenic lines with strong or intermediate phenotypes is indicated by the numbers in parentheses below the flower images. The strong RNAi lines of *PELAN*, *MlANbHLH1* and *MlWD40a* accurately reproduce the phenotypes of *boo3*, *boo14* and *boo5*, respectively, including the stem base color and seed coat color (not shown). Note that no EMS-induced mutants are available for *NEGAN*. (e) Quantitative RT-PCR of the four genes at the 15-mm corolla stage. *MlUBC* was used as the reference gene. All four genes show substantial knock-down in their corresponding RNAi lines. Asterisks highlight the complete absence of *NEGAN* transcripts in the *NEGAN* and *MlWD40a* RNAi lines after 40 cycles of PCR. WT, wild-type. Bars, ± 1 SD from three biological replicates.

NEGAN, but not PELAN, is involved in a self-activation loop

In order to verify that the target regulatory genes have been specifically knocked down in the RNAi lines, we selected three strong lines of each gene for molecular characterization. We first verified transgene presence in these RNAi lines by PCR using transgene specific primers (Fig. S5), then performed semi-quantitative RT-PCR (Fig. S5) and qRT-PCR (Fig. 6e) at the 15-mm corolla stage to examine gene expression. These two sets of experiments consistently showed that in the *PELAN* RNAi lines, the
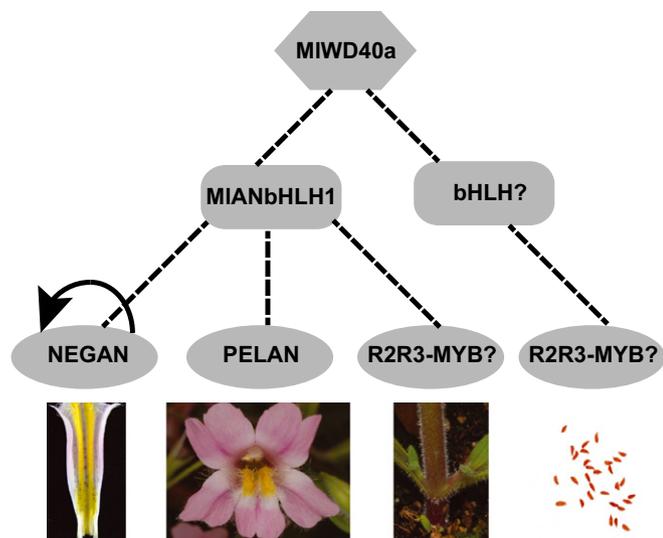
Fig. 7 A regulatory network model of anthocyanin pigmentation in *Mimulus lewisii* LF10. The bottom plant images (from left to right) show the color of nectar guide spots, petal lobes, stem base and seed coat, respectively. The dashed lines indicate putative protein–protein interactions; the arrow indicates self-activation.

expression level of *PELAN* is substantially lower than the wild-type control, and the expression levels of *NEGAN*, *MlANbHLH1* and *MlWD40a* are not affected (Figs 6e,S5). Similarly, in the *MlANbHLH1* RNAi lines, *MlANbHLH1* shows clear downregulation but the other genes are not affected (except *NEGAN*, see the next paragraph).

Interestingly, in the *NEGAN* RNAi lines, while the other genes remain unaffected as expected, the *NEGAN* gene itself showed 100% knock-down (no *NEGAN* transcripts were detected after 40 cycles of PCR; Fig. 6e). This is somewhat surprising because based on our experience with several transcription factor genes characterized in the LF10 background (Yuan *et al.*, 2013a,b; this study), RNAi usually results in 70–90% knock-down in strong transgenic lines, not a complete knock-down. Even more intriguingly, in the *MlWD40a* RNAi lines, not only was *MlWD40a* clearly knocked down, but expression of *NEGAN* also becomes undetectable, although the expression of *PELAN* is not affected (Fig. 6e). This suggests that MlWD40a is required for *NEGAN*, but not *PELAN*, expression. Assuming that MlWD40a functions as part of the MBW complex activating *NEGAN* expression, we could infer that *NEGAN* is involved in an autoregulatory feed-back loop, which explains its 100% knock-down in the *NEGAN* RNAi lines. If *NEGAN* is indeed activated by the 'NEGAN-MlANbHLH1-MlWD40a' complex itself, one would predict that *MlANbHLH1* RNAi should also result in substantial down-regulation of *NEGAN*. However, our qPCR experiment showed only c. 50% *NEGAN* knock-down (Fig. 6e). This is probably due to a redundant *bHLH* factor (e.g. *MlANbHLH2*) expressed in the nectar guide, consistent with the 'weak anthocyanin spot' pheno-type of *boo14* and the *MlANbHLH1* RNAi lines.

In order to further test the *NEGAN* self-activation model, we generated 35S:NEGAN over-expression lines using the full-length *NEGAN* CDS without 5′ or 3′ UTR sequences. We focused our analysis on the leaf tissue, which expresses

MlANbHLH1 and *MlWD40a* at comparable levels to flowers, but expresses no subgroup 6 *R2R3-MYBs* (unpublished leaf transcriptome data). The latter feature is important because that means any detectable endogenous *NEGAN* transcript, which can be distinguished from the transgene transcript by the presence of UTRs, must be activated by the transgene. We obtained 50 inde-pendent 35S:NEGAN lines. A representative whole plant pheno-type is shown in Fig. 8(a). We selected three 35S:NEGAN lines with strong leaf phenotypes (an example shown in Fig. 8b), and performed RT-PCR using primers that can amplify part of the 3′UTR of *NEGAN*. Our results clearly showed that the endoge-nous *NEGAN* gene can be activated by the *NEGAN* transgene in the over-expression lines, and no *NEGAN* expression was detect-able in the wild-type controls (Fig. 8c).

Independent losses of *PELAN* expression (via different mechanisms) explain natural yellow-flowered *M. cardinalis* populations

With a basic model of floral anthocyanin regulation established in *M. lewisii* (Fig. 7), next we explore the possibility of using this model to explain some simple cases of flower color variation in other *Mimulus* species. The first case involves two natural *M. cardinalis* populations with yellow flowers.

M. cardinalis is the sister species of *M. lewisii* (Beardsley *et al.*, 2003) and is typically red-flowered (Fig. 9a) due to the combination of high concentrations of both anthocyanins and carotenoids. The two yellow populations are found in the
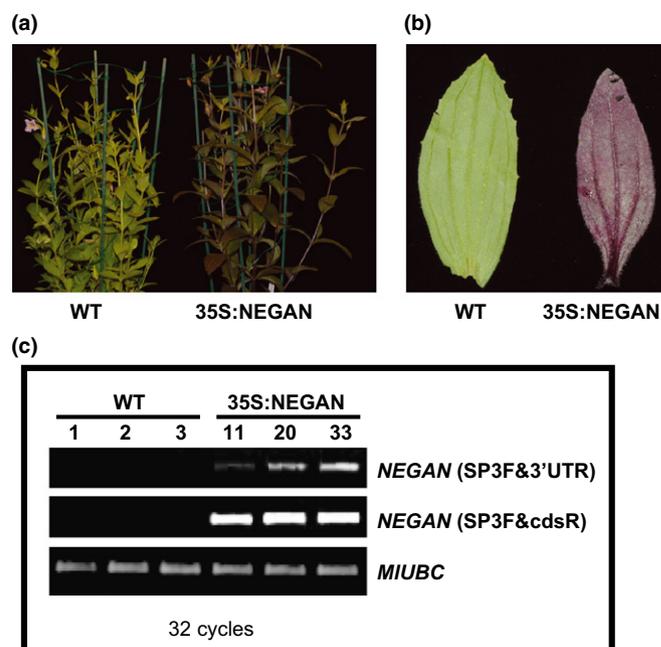
Fig. 8 *NEGAN* transgene activates the endogenous *NEGAN* gene expression in *Mimulus lewisii* over-expression lines. (a) A representative 35S:NEGAN transgenic line showing the whole-plant phenotype. (b) The entire leaf can be changed to dark purple in strong 35S:NEGAN lines. (c) RT-PCR showing activation of the endogenous *NEGAN* gene, which can be distinguished from the transgene by the presence of 3′UTR, in three independent 35S:NEGAN lines. WT, wild-type.

M. cardinalis (CE10)

**M. cardinalis (SM)
(Siskiyou Mts., OR)**

**M. cardinalis (CI)
(Cedros Island, Baja CA)**
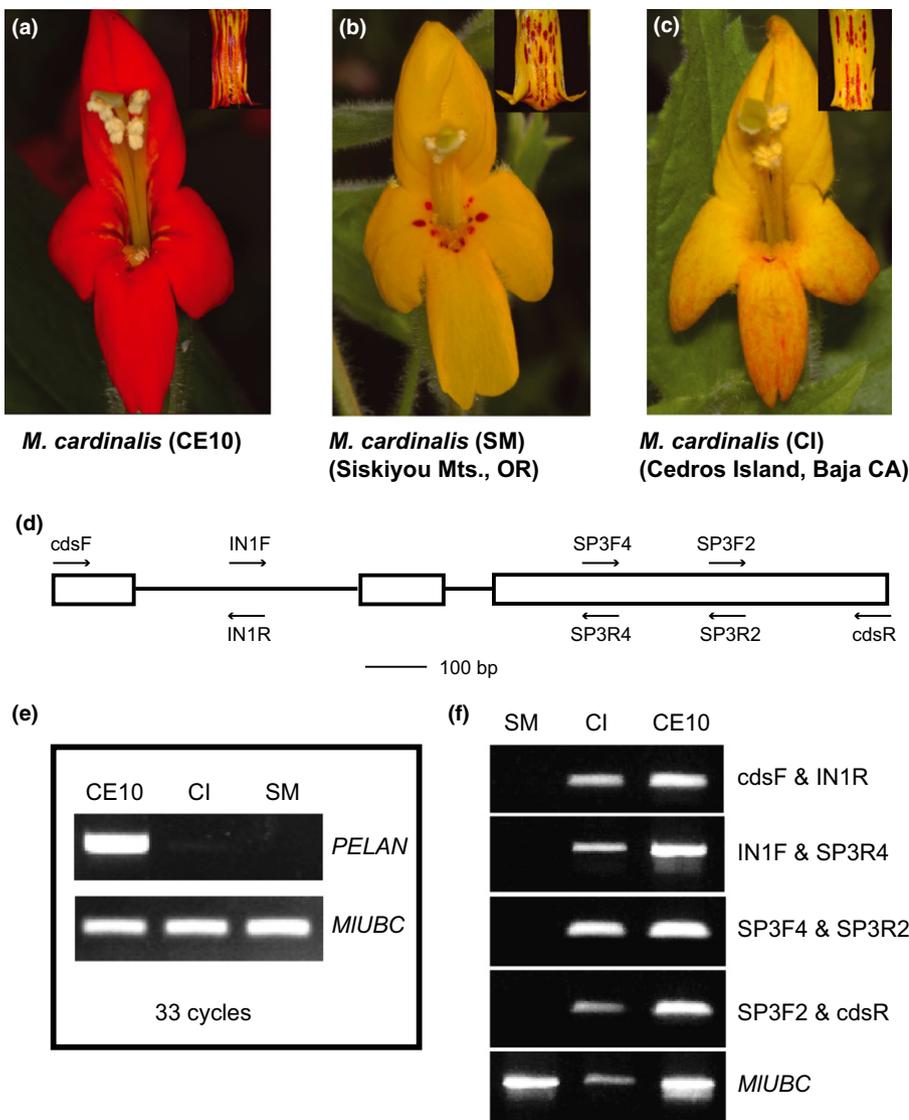
Fig. 9 Molecular characterization of the two natural yellow-flowered *Mimulus cardinalis* populations. (a) *M. cardinalis* inbred line CE10 showing the typical red flower. (b) *M. cardinalis* (SM). (c) *M. cardinalis* (CI). The dissected nectar guide on the upper-right corner of the flower images show the anthocyanin spots in the nectar guide. (d) Structure of the *PELAN* gene and the positions of PCR primers used in RT-PCR and genomic PCR. (e) RT-PCR (primer pair SP3F4 & SP3R2) shows no expression of *PELAN* in CI or SM. (f) Genomic PCR by multiple primer pairs suggest that *PELAN* has probably been deleted from the SM genome. Primer sequences are listed in Table S4 (also see Fig. S7). GenBank accession numbers for the *M. cardinalis* CE10 and CI *PELAN* gene sequences are KJ595587–KJ595588.

northern (Siskiyou Mountains, Oregon) and southern (Cedros Island, Baja California) margins of the *M. cardinalis* geographic range (Vickery, 1995; Paul *et al.*, 2011), and hereafter will be referred to as *M. cardinalis* SM and *M. cardinalis* CI, respectively. Both *M. cardinalis* SM and CI lack anthocyanins in the petal lobe, but have strong anthocyanin pigmentation in the nectar guide (Fig. 9b,c). Genetic crosses with *M. cardinalis* CE10 (Fig. 9a) show that yellow is recessive to red. The similarity of floral anthocyanin pigmentation pattern between *M. cardinalis* SM, CI and the *M. lewisii boo3* mutant suggests that the phenotype of all three is likely to be caused by the loss of function of the same gene, *PELAN*. Indeed, complementation crosses between the three suggest that they are all allelic (Fig. S6).

Next we ask: What is the molecular nature of the loss-of-function *pelan* allele of *M. cardinalis* SM and CI? These two populations are geographically isolated from each other and display slightly different phenotypes (Fig. 9b,c), which suggests that they have evolved independently from the typical red-flowered

phenotype. To address this question, we first attempted to amplify and sequence the full-length CDS of *PELAN* from *M. cardinalis* SM and CI corolla cDNA at the stage corresponding to the 15-mm corolla stage of *M. lewisii* LF10, to examine potential coding DNA mutations. However, although the *PELAN* CDS could be readily amplified from both *M. lewisii* LF10 and the red-flowered *M. cardinalis* CE10 corolla cDNA, it could not be amplified from either *M. cardinalis* SM or CI. This suggests that *PELAN* may not be expressed in the yellow *M. cardinalis*. Further RT-PCR experiments corroborate this inference (Fig. 9e). To rule out the possibility that this failure of detecting *PELAN* expression by RT-PCR is due to primer mismatch in SM and CI, we performed PCR on the genomic DNA as a control. The RT-PCR primers produced a band of the expected size with the CI genomic template, but, surprisingly, still failed to amplify any products in SM (Fig. 9f). These results indicate an intriguing possibility that the *PELAN* gene has been deleted from the SM genome. To test this idea, we designed multiple pairs of primers, from both exonic and intronic regions

and with sequences conserved between LF10, CE10 and CI (Figs 9d,S7) – all of these primer pairs produced a clean band of expected size in CE10 and CI, but no bands in SM (Fig. 9f).

Taken together, these results suggest that the losses of petal lobe anthocyanins in *M. cardinalis* SM and CI are caused by independent molecular lesions in the *PELAN MYB* gene. The loss of *PELAN* expression in SM is linked with the probable deletion of the entire gene in the SM genome, whereas the loss of *PELAN* expression in CI is most likely to be caused by *cis*-regulatory changes because the complementation crosses have ruled out *trans*-acting factors as a potential cause.

The *NEGAN* ortholog is the only anthocyanin-activating *MYB* expressed in the *M. guttatus* corolla

The second case of using the *M. lewisii* model to explain floral anthocyanin pigmentation patterns in other *Mimulus* involves *M. guttatus*, for which many genomic resources are available (Wu *et al.*, 2008; Hellsten *et al.*, 2013). *M. guttatus* belongs to the 'yellow flower' clade (the top clade in Fig. 1), where most species do not have petal lobe anthocyanins but do have anthocyanin spots in the corolla throat (Fig. 10a), presumably serving as a nectar guide for pollinators.

The *M. guttatus* genome has five *R2R3-MYB* genes (*MgMYB1-5*) that are classified in the anthocyanin-activating group (Cooley *et al.*, 2011; Fig. 2a). Phylogenetic analysis suggests that *MgMYB1-3* groups together with *PELAN* and *MgMYB5* groups with *NEGAN*. *MgMYB4* represents a more divergent lineage (Fig. 2a). Considering that *NEGAN* is the *MYB* that controls anthocyanin spot formation in the *M. lewisii* nectar guide, we predicted that the *NEGAN* ortholog, *MgMYB5*, is the only anthocyanin-activating *MYB* gene expressed in the *M. guttatus* corolla, to explain the fact that anthocyanins are restricted to the nectar guide spots in the *M. guttatus* flower. RT-PCR experiments clearly show that this is the case (Fig. 10b). These results suggest that the function of *NEGAN* – making anthocyanin spots in the nectar guide – is conserved between these two major clades of *Mimulus* (Fig. 1), and our *M. lewisii* model can be used to explain anthocyanin pigmentation patterns even in distantly related species.

Discussion

In this study we have identified four major transcriptional regulators of anthocyanin pigmentation in *M. lewisii* flowers, including the *MlWD40a* and *MlANbHLH1* genes that control anthocyanin pigmentation in the entire corolla, the *PELAN R2R3-MYB* that controls the petal lobe color, and the *NEGAN R2R3-MYB* that regulates the anthocyanin spot formation in the nectar guide. Furthermore, through two case studies in *M. cardinalis* and *M. guttatus*, we have demonstrated that the establishment of a baseline floral anthocyanin regulation model in *M. lewisii* is of great value towards understanding the molecular bases underlying the astonishing diversity of floral anthocyanin pigmentations patterns in other *Mimulus* species.

The identities and functions of all four regulatory genes are supported by multiple lines of evidence. Overall the different degrees of pleiotropy of these genes are consistent with the MBW regulatory network model characterized in other plant species (Zhang *et al.*, 2003; Koes *et al.*, 2005; Hichri *et al.*, 2011) – *WD40* is the most pleiotropic, while the *R2R3-MYB* is the least pleiotropic and tends to act in a tissue-specific fashion. The identity and function of *PELAN* are supported by multiple alleles from both EMS-induced (*boo3*) and natural mutants (*M. cardinalis* SM and CI), and by gene-specific RNAi knock-down experiments. The identity and function of *NEGAN* are supported by the tissue-specific expression pattern (only in the nectar guide) and the distinct RNAi transgenic phenotype, with only the anthocyanin spots in the nectar guide being affected. Similarly, a combination of EMS-induced loss-of-function alleles and RNAi phenocopies confirm the identity and function of *MlANbHLH1* and *MlWD40a*.

The autoregulation of *NEGAN* is noteworthy. Although the regulatory role of the MBW complex in ABP structural gene expression is highly conserved across all flowering plants characterized to date (Koes *et al.*, 2005; Glover, 2007; Hichri *et al.*, 2011; Davies *et al.*, 2012), the transcriptional regulation of the MBW regulatory genes themselves is less understood. In *Petunia*, *Arabidopsis* and grapevine (*Vitis vinifera*), some anthocyanin-activating *bHLH* genes can be activated by the MBW complex itself (Spelt *et al.*, 2000; Baudry *et al.*, 2006; Hichri *et al.*,

(a) **(b)**

Fig. 10 Anthocyanin pigmentation in *Mimulus guttatus* inbred line IM767. (a) Flower images showing the anthocyanin spots in the nectar guide. (b) The *NEGAN* ortholog is the only anthocyanin-activating *MYB* expressed in the corolla of IM767, indicated by the red arrow. Genomic DNA was used as control to test primer quality. The *M. guttatus UBC* ortholog (*MgUBC*) was used as a reference gene. The larger size of the *MgUBC* genomic amplicon is due to the presence of an intron in the amplified fragment. Primer sequences are listed in Table S1.

2010). However, in maize, the three components of the MBW complex seem to be independently regulated (Carey *et al.*, 2004). The *bHLH* in *M. lewisii* (*MlANbHLH1*) does not seem to be regulated by the MBW complex, either, as neither *MlWD40a* nor *PELAN/NEGAN* knock-down affects *MlANbHLH1* expression. Little is known about the transcriptional control of the *R2R3-MYBs*, which usually show more tissue-specific expression patterns (e.g. *PELAN* and *NEGAN*). One interesting exception is the apple *MYB10* that activates anthocyanin biosynthesis in the red-fleshed apple varieties. The rearrangement of the *MYB10* upstream regulatory region in some apple varieties results in a tandem repeat of a 23-bp sequence that serves as binding site for the MYB10 protein itself, leading to the autoregulation of *MYB10* (Espley *et al.*, 2009). Our RNAi and over-expression experiments strongly suggest that *NEGAN*, but not *PELAN*, is activated by the MBW complex itself (Figs 6e, 8). The fact that the MBW complex regulates *NEGAN* but not *PELAN* suggests that these two *R2R3-MYB* genes have evolved distinct *cis*-elements that respond to different transcriptional regulators.

It is also interesting to note the link between the self-activation of *NEGAN* and the formation of anthocyanin spots rather than a solid pattern in the nectar guide. Formation of spotty patterns in biological objects is often explained by the reaction-diffusion model (Turing, 1952) or various modified versions of this model (Meinhardt, 1982; Meinhardt & Gierer, 2000; Kondo & Miura, 2010). The essence of these reaction-diffusion based models is an interacting network that contains a local autocatalytic feedback loop and a long-range inhibitory feedback loop (Meinhardt & Gierer, 2000). The activation of *NEGAN* by the NEGAN-MlANbHLH1-WD40a complex itself forms such an autocatalytic feedback, although what may constitute the long-range inhibitory feedback is not yet clear. Notably, the reaction-diffusion model has previously been proposed as a potential mechanism generating multicellular pigmented petal spots by Davies *et al.* (2012). In fact, the transcriptional network regulating anthocyanin pigmentation in *Petunia hybrida* contains both an autocatalytic activator and a potential long-range repressor – the *bHLH* gene *AN1* is activated by the MBW complex itself, which also activates an *R3-MYB* repressor, *MYBx*; MYBx inhibits the activity of AN1 and is capable of intercellular movement (Albert *et al.*, 2014). However, the flowers of *Petunia hybrida* do not usually display anthocyanin spots. This suggests that the existence of such feedback loops may be required, but may not necessarily be sufficient for spot formation. The specific properties of each component (e.g. the relative diffusion rate of the long-range inhibitor to the short-range activators) could be critical to form different patterns (Kondo & Miura, 2010). It will be interesting to determine whether one can 'engineer' anthocyanin spots in *Petunia* petals by fine-tuning the properties of these activators and repressors through transgenic manipulations.

In addition to the four major transcriptional activators, a single-repeat *R3-MYB* gene, *ROI1*, similar to the *Petunia MYBx*, has been previously identified as a negative regulator of anthocyanin biosynthesis in *M. lewisii* petal lobes (Yuan *et al.*, 2013a). Yeast-two-hybrid experiments suggest that PELAN, NEGAN and ROI1 can all interact with MlANbHLH1 in yeast (Y-W. Yuan &

H.D. Bradshaw, unpublished data), supporting the hypothesis that ROI1 negatively regulates anthocyanin biosynthesis by competing with the R2R3-MYB activators for the limited supply of bHLH proteins. At first glance ROI1 may appear as a good candidate fulfilling the 'inhibitory feedback' requirement as a competitor of NEGAN, to explain the spotty anthocyanin pattern in the nectar guide. However, this is unlikely to be the case because *ROI1* has very low expression level in the nectar guide compared to petal lobes. Consistent with this, knocking down *ROI1* in *M. lewisii* does not seem to have any effects on the nectar guide anthocyanin spots (Yuan *et al.*, 2013a). There probably exists another yet-to-be-identified repressor specifically expressed in the nectar guide, forming inhibitory interactions with *NEGAN*.

The successful prediction of the molecular bases underlying the natural yellow *M. cardinalis* populations and the *M. guttatus* floral anthocyanin patterns using our *M. lewisii* model is particularly satisfying. The two yellow *M. cardinalis* populations are found in the northern (Siskiyou Mountains of Oregon) and southern limit (Cedros Island, Baja California) of the species range (Vickery, 1995; Paul *et al.*, 2011), respectively, and thus represent a classical example of new forms evolving from isolated populations on the periphery of a species range (Mayr, 1976). Pollination observations in a common garden environment using the typical red flowers and the mutant yellow forms suggested that the color change from red to yellow in *M. cardinalis* is probably sufficient to 'initiate partial, incipient reproductive isolation' (Vickery, 1995). As such, by revealing the loss of *PELAN* expression as the cause for this color change, we might have recovered the molecular bases of two parallel incipient speciation events.

The *M. cardinalis* results also provide further evidence supporting the notion that phenotypic variation within and between species often involve the same 'hotspot' genes (Stern, 2011; Streisfeld & Rausher, 2011). The petal lobe-specific *R2R3-MYB*, *PELAN*, is such a hotspot gene occupying 'a privileged position' (Stern, 2011) in a genetic network that can 'maximize the phenotypic output' by causing the downregulation of multiple anthocyanin biosynthetic genes, but meanwhile can 'minimize pleiotropy' by changing only the petal lobe color without affecting any other traits. However, there could be many different ways leading to loss-of-function of the same hotspot gene. In the two isolated *M. cardinalis* populations, loss of *PELAN* expression has evolved independently by two completely different mechanisms: one is probably caused by *cis*-regulatory changes and the other likely involves the deletion of the entire gene.

The *M. guttatus* case study demonstrates that our *M. lewisii* floral anthocyanin regulation model is applicable not only to closely related species, but also to relatively distantly related species (for phylogenetic relationships see Fig. 1). None of the *R2R3-MYBs* in the *PELAN* clade (Fig. 2a) is expressed in *M. guttatus* corolla (Fig. 10b), explaining the lack of anthocyanins in petal lobes. The *NEGAN* orthog, *MgMYB5*, appears to be responsible for the formation of nectar guide anthocyanin spots in *M. guttatus*.

Ultimately, we would like to extend the two simple case studies presented here and use the *M. lewisii* model to examine the molecular tinkering underlying the evolution of more complex

floral anthocyanin patterns, such as the different anthocyanin shades in the petal lobes of *M. pulchellus* or the novel 'spider-web' pattern in *M. pictus* (Fig. 1). With the necessary mutant lines and the powerful transgenic tools available in *M. lewisii*, we can introduce genomic copies (including upstream regulatory regions) of each major anthocyanin regulator from other species, on a gene-by-gene basis, into the corresponding *M. lewisii* mutant background. In doing so, we will be able to address the question of how many gene replacements are required to change one complex pigmentation pattern into another. More importantly, given the gradient of phylogenetic distances between the various species and *M. lewisii*, the gene-by-gene replacement strategy will allow us to test whether phenotypic evolution at deeper phylogenetic levels (between distantly related species) involves different kinds of mechanisms, genes or mutations to those involved with phenotypic evolution within species or between closely related species (Stern & Orgogozo, 2009).

References

Albert NW, Davies KM, Lewis DH, Zhang HB, Montefiori M, Brendolise C, Boase MR, Ngo H, Jameson PE, Schwinn KE. 2014. A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in Eudicots. *Plant Cell* 26: 962–980.

Albert NW, Lewis DH, Zhang H, Schwinn KE, Jameson PE, Davies KM. 2011. Members of an R2R3-MYB transcription factor family in *Petunia* are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. *Plant Journal* 65: 771–784.

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25: 3389–3402.

Baudry A, Caboche M, Lepiniec L. 2006. TT8 controls its own expression in a feedback regulation involving TTG1 and homologous MYB and bHLH factors, allowing a strong and cell-specific accumulation of flavonoids in *Arabidopsis thaliana*. *Plant Journal* 46: 768–779.

Beardsley PM, Schoenig SE, Whittall JB, Olmstead RG. 2004. Patterns of evolution in western North American *Mimulus* (Phrymaceae). *American Journal of Botany* 91: 474–489.

Beardsley PM, Yen A, Olmstead RG. 2003. AFLP phylogeny of *Mimulus* section *Erythranthe* and the evolution of hummingbird pollination. *Evolution* 57: 1397–1410.

Borevitz JO, Xia YJ, Blount J, Dixon RA, Lamb C. 2000. Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12: 2383–2393.

Bradshaw HD, Schemske DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* 426: 176–178.

Carey CC, Strahle JT, Selinger DA, Chandler VL. 2004. Mutations in the *pale aleurone color1* regulatory gene of the *Zea mays* anthocyanin pathway have distinct phenotypes relative to the functionally similar *TRANSPARENT TESTA GLABRA1* gene in *Arabidopsis thaliana*. *Plant Cell* 16: 450–464.

Cooley AM, Modliszewski JL, Rommel ML, Willis JH. 2011. Gene duplication in *Mimulus* underlies parallel floral evolution via independent *trans*-regulatory changes. *Current Biology* 21: 700–704.

Cooley AM, Willis JH. 2009. Genetic divergence causes parallel evolution of flower color in Chilean *Mimulus*. *New Phytologist* 183: 729–739.

Davies KM, Albert NW, Schwinn KE. 2012. From landing lights to mimicry: the molecular regulation of flower colouration and mechanisms for pigmentation patterning. *Functional Plant Biology* 39: 619–638.

Des Marais DL, Rausher MD. 2010. Parallel evolution at multiple levels in the origin of hummingbird pollinated flowers in *Ipomoea*. *Evolution* 64: 2044–2054.

Earley KW, Haag JR, Pontes O, Opper K, Juehne T, Song KM, Pikaard CS. 2006. Gateway-compatible vectors for plant functional genomics and proteomics. *Plant Journal* 45: 616–629.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* 32: 1792–1797.

Espley RV, Brendolise C, Chagne D, Kutty-Amma S, Green S, Volz R, Putterill J, Schouten HJ, Gardiner SE, Hellens RP et al. 2009. Multiple repeats of a promoter segment causes transcription factor autoregulation in red apples. *Plant Cell* 21: 168–183.

Feller A, Machemer K, Braun EL, Grotewold E. 2011. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant Journal* 66: 94–116.

Glover BJ. 2007. *Understanding flowers and flowering: an integrated approach.* Oxford, UK: Oxford University Press.

Goodrich J, Carpenter R, Coen ES. 1992. A common gene regulates pigmentation pattern in diverse plant species. *Cell* 68: 955–964.

Grossenbacher DL, Whittall JB. 2011. Increased floral divergence in sympatric monkeyflowers. *Evolution* 65: 2712–2718.

Grotewold E. 2006. The genetics and biochemistry of floral pigments. *Annual Review of Plant Biology* 57: 761–780.

Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC. 2003. The basic helix-loop-helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. *Molecular Biology and Evolution* 20: 735–747.

Hellsten U, Wright KM, Jenkins J, Shu S, Yuan YW, Wessler SR, Schmutz J, Willis JH, Rokhsar DS. 2013. Fine-scale variation in meiotic recombination in *Mimulus* inferred from population shotgun sequencing. *Proceedings of the National Academy of Sciences, USA* 110: 19 478–19 482.

Hichri I, Barrieu F, Bogs J, Kappel C, Delrot S, Lauvergeat V. 2011. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. *Journal of Experimental Botany* 62: 2465–2483.

Hichri I, Heppel SC, Pillet J, Leon C, Czemmel S, Delrot S, Lauvergeat V, Bogs J. 2010. The basic helix-loop-helix transcription factor MYC1 is involved in the regulation of the flavonoid biosynthesis pathway in grapevine. *Molecular Plant* 3: 509–523.

Hiesey W, Nobs MA, Björkman O. 1971. *Experimental studies on the nature of species. V. Biosystematics, genetics, and physiological ecology of the Erythranthe section of Mimulus.* Washington, DC, USA: Carnegie Institute of Washington publ. no. 628.

Hoballah ME, Gubitz T, Stuurman J, Broger L, Barone M, Mandel T, Dell'Olivo A, Arnold M, Kuhlemeier C. 2007. Single gene-mediated shift in pollinator attraction in *Petunia*. *Plant Cell* 19: 779–790.

Hopkins R, Rausher MD. 2011. Identification of two genes causing reinforcement in the Texas wildflower *Phlox drummondii*. *Nature* 469: 411–414.

Koes R, Verweij W, Quattrocchio F. 2005. Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science* 10: 236–242.

Kondo S, Miura T. 2010. Reaction-diffusion model as a framework for understanding biological pattern formation. *Science* 329: 1616–1620.

Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M. 2006. Genetics and biochemistry of seed flavonoids. *Annual Review of Plant Biology* 57: 405–430.

Ludwig SR, Habera LF, Dellaporta SL, Wessler SR. 1989. Lc, a member of the maize R-gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcriptional activators and contains the Myc-homology region. *Proceedings of the National Academy of Sciences, USA* 86: 7092–7096.

Ma HM, Pooler M, Griesbach R. 2009. Anthocyanin regulatory/structural gene expression in *Phalaenopsis*. *Journal of the American Society for Horticultural Science* 134: 88–96.

Martin C, Prescott A, Mackay S, Bartlett J, Vrijlandt E. 1991. Control of anthocyanin biosynthesis in flowers of *Antirrhinum majus*. *Plant Journal* 1: 37–49.

Martins TR, Berg JJ, Blinka S, Rausher MD, Baum DA. 2013. Precise spatio-temporal regulation of the anthocyanin biosynthetic pathway leads to petal spot formation in *Clarkia gracilis* (Onagraceae). *New Phytologist* 197: 958–969.

Mayr E. 1976. *Evolution and the diversity of life*. Cambridge, MA, USA: Harvard University Press.

Meinhardt H. 1982. *Models of biological pattern formation*. London, UK: Academic Press.

Meinhardt H, Gierer A. 2000. Pattern formation by local self-activation and lateral inhibition. *BioEssays* 22: 753–760.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5: 621–628.

Nakatsuka T, Yamada E, Saito M, Fujita K, Nishihara M. 2013. Heterologous expression of gentian MYB1R transcription factors suppresses anthocyanin pigmentation in tobacco flowers. *Plant Cell Reports* 32: 1925–1937.

Owen CR, Bradshaw HD. 2011. Induced mutations affecting pollinator choice in *Mimulus lewisii* (Phrymaceae). *Arthropod-Plant Interactions* 5: 235–244.

Paul JR, Sheth SN, Angert AL. 2011. Quantifying the impact of gene flow on phenotype-environment mismatch: a demonstration with the scarlet monkeyflower *Mimulus cardinalis*. *American Naturalist* 178: S62–S79.

Paz-Ares J, Ghosal D, Wienand U, Peterson PA, Saedler H. 1987. The regulatory C1 locus of *Zea mays* encodes a protein with homology to MYB protooncogene products and with structural similarities to transcriptional activators. *EMBO Journal* 6: 3553–3558.

Pires N, Dolan L. 2010. Origin and diversification of basic-helix-loop-helix proteins in plants. *Molecular Biology and Evolution* 27: 862–874.

Quattrocchio F, Wing JF, van der Woude K, Mol JNM, Koes R. 1998. Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. *Plant Journal* 13: 475–488.

Quattrocchio F, Wing J, van der Woude K, Souer E, de Vetten N, Mol J, Koes R. 1999. Molecular analysis of the *anthocyanin2* gene of petunia and its role in the evolution of flower color. *Plant Cell* 11: 1433–1444.

Ruiz-Sola M, Rodriguez-Concepcion M. 2012. Carotenoid biosynthesis in *Arabidopsis*: a colorful pathway. *The Arabidopsis Book* 10: e0158.

Schwinn K, Venail J, Shang YJ, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C. 2006. A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *Plant Cell* 18: 831–851.

Shang YJ, Venail J, Mackay S, Bailey PC, Schwinn KE, Jameson PE, Martin CR, Davies KM. 2011. The molecular basis for venation patterning of pigmentation and its effect on pollinator attraction in flowers of *Antirrhinum*. *New Phytologist* 189: 602–615.

Smith SD, Rausher MD. 2011. Gene loss and parallel evolution contribute to species difference in flower color. *Molecular Biology and Evolution* 28: 2799–2810.

Sobel J, Streisfeld M. 2013. Flower color as a model system for studies of plant evo-devo. *Frontiers in Plant Science* 4: Article 321.

Spelt C, Quattrocchio F, Mol JNM, Koes R. 2000. *anthocyanin1* of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell* 12: 1619–1631.

Stamatakis A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690.

Stern DL. 2011. *Evolution, development, and the predictable genome*. Greenwood Village, CO, USA: Roberts and Co.

Stern DL, Orgogozo V. 2009. Is genetic evolution predictable? *Science* 323: 746–751.

Stracke R, Werber M, Weisshaar B. 2001. The *R2R3-MYB* gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* 4: 447–456.

Streisfeld MA, Kohn JR. 2005. Contrasting patterns of floral and molecular variation across a cline in *Mimulus aurantiacus*. *Evolution* 59: 2548–2559.

Streisfeld MA, Rausher MD. 2011. Population genetics, pleiotropy, and the preferential fixation of mutations during adptive evolution. *Evolution* 65: 629–642.

Streisfeld MA, Young WN, Sobel JM. 2013. Divergent selection drives genetic differentiation in an *R2R3-MYB* transcription factor that contributes to incipient speciation in *Mimulus aurantiacus*. *PLoS Genetics* 9: e1003385.

Talavera G, Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* 56: 564–577.

Turing AM. 1952. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 237: 37–72.

de Vetten N, Quattrocchio F, Mol J, Koes R. 1997. The *an11* locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. *Genes & Development* 11: 1422–1434.

Vickery RK. 1995. Speciation in *Mimulus*, or, can a simple flower color mutant lead to species divergence. *Great Basin Naturalist* 55: 177–180.

Vickery RK, Olson RL. 1956. Flower color inheritance in the *Mimulus cardinalis* complex. *Journal of Heredity* 47: 195–199.

Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundell TL, Esch JJ, Marks MD, Gray JC. 1999. The *TRANSPARENT TESTA GLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. *Plant Cell* 11: 1337–1349.

Wu CA, Lowry DB, Cooley AM, Wright KM, Lee YW, Willis JH. 2008. *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity* 100: 220–230.

Yamagishi M, Shimoyamada Y, Nakatsuka T, Masuda K. 2010. Two *R2R3-MYB* genes, homologs of *Petunia AN2*, regulate anthocyanin biosyntheses in flower tepals, tepal spots and leaves of Asiatic hybrid lily. *Plant and Cell Physiology* 51: 463–474.

Yamagishi M, Toda S, Tasaki K. 2014. The novel allele of the *LhMYB12* gene is involved in splatter-type spot formation on the flower tepals of Asiatic hybrid lilies (*Lilium* spp.). *New Phytologist* 201: 1009–1020.

Yuan YW, Sagawa JM, Di Stilio VS, Bradshaw HD. 2013b. Bulk segregant analysis of an induced floral mutant identifies a *MIXTA*-like *R2R3 MYB* controlling nectar guide formation in *Mimulus lewisii*. *Genetics* 194: 523–528.

Yuan YW, Sagawa JM, Young RC, Christensen BJ, Bradshaw HD. 2013a. Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. *Genetics* 194: 255–263.

Zhang F, Gonzalez A, Zhao MZ, Payne CT, Lloyd A. 2003. A network of redundant bHLH proteins functions in all TTG1-dependent pathways of *Arabidopsis*. *Development* 130: 4859–4869.

Zhu HF, Fitzsimmons K, Khandelwal A, Kranz RG. 2009. CPC, a single-repeat R3 MYB, is a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. *Molecular Plant* 2: 790–802.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 RT-PCR of different paralogs of *CHS*, *F3H*, and the *MBW* regulatory genes.

Fig. S2 Alignment of subgroup 6 R2R3-MYBs.

Fig. S3 Alignment of subgroup IIIf bHLHs.

Fig. S4 Alignment of WD40 proteins.

Fig. S5 Molecular characterization of the RNAi transgenic lines.

Fig. S6 Complementation test between *Mimulus lewisii boo3*, *M. cardinalis* SM and *M. cardinalis* CI.

Fig. S7 Alignment of *PELAN* sequences from *Mimulus lewisii* LF10, *M. cardinalis* CE10 and *M. cardinalis* CI.

Table S1 Primers used in RT-PCR experiments

Table S2 Primers used for candidate gene sequencing of the *boo* mutants

Table S3 Primers used for constructing RNAi and over-expression plasmids

Table S4 Primers used for *PELAN* amplification in *Mimulus cardinalis* strains

Please note: Wiley Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.

About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews.

- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <25 days. There are **no page or colour charges** and a PDF version will be provided for each article.

- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.

- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)

- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**