

## Supplementary Material

Article title: Ancient gene duplications, rather than polyploidization, facilitate diversification of petal pigmentation patterns in *Clarkia gracilis* (Onagraceae)

Authors: Rong-Chien Lin, Mark D. Rausher

**Table S1.** Primers used in this study.

**Table S2.** Anthocyanin genes identified from the transcriptomes of *Clarkia gracilis albicaulis*.

**Table S3.** Synonymous substitution rate ( $d_s$ ) along the branches leading to *Clarkia gracilis* and to its progenitor species, *C. amoena huntiana* and *C. lassenensis*.

**Table S4.** Voucher information for the *Clarkia* plants used in this study.

**Fig. S1.** Alignment of *Clarkia* R2R3-MYB sequences (A) *MYB1*, (B) *MYB6*, (C) *MYB11* and (D) *MYB12* from *C. gracilis sonomensis*, *C. g. albicaulis*, *C. amoena huntiana* and *C. lassenensis*.

**Fig. S2.** Spatiotemporal expression of *MYB1*, *MYB6*, *MYB11* and *MYB12* in the petals of (A) *Clarkia amoena huntiana*, (B) *C. lassenensis*, (C) pink-cupped *C. gracilis sonomensis*, (D) white-cupped *C. g. sonomensis*, (E) *C. g. albicaulis*.

**Fig. S3.** Gene tree used to estimate  $d_s$  (synonymous substitution rate) along branches.

**Methods S1.** Transcriptomics

Table S1. Primers used in this study. <sup>†</sup>Primers from Martins *et al.* 2017. <sup>‡</sup>Primers from Lin and Rausher 2021. <sup>§</sup>Semi-qPCR primers used for *Clarkia lassenensis*.

---

**Primers for coding region sequencing**

---

Target species: *Clarkia amoena huntiana*

---

Primer Name	Primer Sequence (5'-3')
cMYB1-F <sup>†</sup>	ATGAATAAGGTAGGACTTAGAAAGG
cMYB1-R <sup>†</sup>	TTAAAATATGTCATCAAAATAAAGCTCATCC
cMYB6-3F <sup>‡</sup>	TGCTACAGAAAGTCTAACGT
cMYB6-1R <sup>‡</sup>	ACCGCTGATTTATTTGAAACCCT
cMYB11-5F <sup>‡</sup>	AAAAACCAGAAGAAAACCCA
cMYB11-6R	AATTTACAGTTTAATCATTTG
cMYB12-9F	AATGAAGGAAGGTCTAAGGAA
cMYB12-5R	CTACAAC TGAAATATGTCGTGATC

---

Target species: *Clarkia lassenensis*

---

Primer Name	Primer Sequence (5'-3')
cMYB6-4F	CATGGGTGGTGTTTCCTTGGA
cMYB6-5R	TTACAGAGAGTTATTCCACAGATCG
cMYB11-7F	AATGAAGGGAGAATTAAGGAAG
cMYB11-5R	TACTTGGAATATGATTCTACACA
cMYB12-11F	ATGCACGCTCTACAATAAAACG
cMYB12-6R	TTTAGTATAACGGCATAATTTCT

---

Target species: *Clarkia gracilis albicaulis*

---

Primer Name	Primer Sequence (5'-3')
cMYB6-3F <sup>‡</sup>	TGCTACAGAAAGTCTAACGT
cMYB6-1R <sup>‡</sup>	ACCGCTGATTTATTTGAAACCCT
cMYB11-3F	GTGAGTCGCATGGCTGT
cMYB11-2R <sup>‡</sup>	CACAGTTTAATCATTTGATTC

---

**Primers for semi-quantitative assessment**

---

Primer Name	Primer Sequence (5'-3')
cMYB1-gF <sup>†</sup>	TGGTCACTCATAGCAGGAAGA
cMYB1-Q1R	TCATCAACCAGTCCGACCAA

CIMYB1-1R <sup>a</sup>	GGGTTATACCCACTCTCAGTTCC
cMYB6-Q1F <sup>‡</sup>	GACGAACCCACCGAATCAAG
cMYB6-Q1R <sup>‡</sup>	CAGATCGTCCCAGTCCCATT
CIMYB6-Q1F <sup>a</sup>	CGGTGTCGGTCTTGGAGATA
CIMYB6-Q1R <sup>a</sup>	TCATCCCAATCCCATTTTCTGC
cMYB11-Q1F <sup>‡</sup>	GACAAGGAGGTGATGAATTGGT
cMYB11-Q1R <sup>‡</sup>	ATTCTACACATTCATGGAGTCGA
cMYB12-Q1F <sup>‡</sup>	GAGCAAGATTCGGTCAAAGT
cMYB12-1R <sup>‡</sup>	TATTCCGTTACAATGAGGCT
GAPDH-Q1F <sup>‡</sup>	GAGGCATCAGAGACCCACAT
GAPDH-Q1R <sup>‡</sup>	CACGACACGAGCTTCACAAA

---

Table S2. Anthocyanin genes identified from the transcriptomes of *Clarkia gracilis albicaulis*. The transcriptomes were obtained from pink background and white band of the *C. g. albicaulis* petals, following the protocol described in supplementary methods S1. Gene expression levels were estimated as FPKM values by mapping reads to the transcriptome references of pink background and white band, separately.

Contig ID	Gene	Expression (FPKM)	
		Pink background	White band
<i>Pink background reference</i>			
RCL6_27068_c0_g1	<i>Chs</i>	1350.66	1138.03
RCL6_21692_c0_g1	<i>Chi</i>	74.56	26.34
RCL6_23311_c0_g4	<i>F3h</i>	45.94	50.03
RCL6_27159_c0_g1	<i>F3'h</i>	15.98	27.28
RCL6_23784_c0_g1	<i>F3'5'h</i>	649.53	381.47
RCL6_23661_c0_g1	<i>Dfr</i>	757.58	121.86
RCL6_26489_c1_g1	<i>Dfr3</i>	66.31	1140.31
RCL6_26565_c4_g2	<i>Ans</i>	628.81	45.37
RCL6_22269_c2_g4	<i>Uf3gt</i>	216.18	154.25
RCL6_24035_c1_g1	<i>MYB6</i>	858.87	595.14
RCL6_21823_c2_g1	<i>MYB11</i>	121	46.94
RCL6_28257_c0_g1	<i>bHLH1</i>	39.6	34.93
RCL6_20240_c0_g1	<i>bHLH2</i>	4.88	5.04
RCL6_23159_c0_g2	<i>WDR1</i>	20.9	14.74
RCL6_21868_c0_g1	<i>WDR2</i>	14.88	13.33
<i>White band reference</i>			
RCL7_26363_c0_g2	<i>Chs</i>	1372.81	846.27
RCL7_20945_c0_g1	<i>Chi</i>	108.08	25.05
RCL7_20216_c0_g4	<i>F3h</i>	78.11	70.13
RCL7_26446_c0_g1	<i>F3'h</i>	16.34	22.74
RCL7_23161_c0_g1	<i>F3'5'h</i>	639.85	290.18
RCL7_28129_c1_g1	<i>Dfr1</i>	610.03	77.41
RCL7_20076_c4_g1	<i>Dfr3</i>	52.54	705.62
RCL7_23829_c0_g1	<i>Ans</i>	579.14	35.97
RCL7_20317_c4_g2	<i>Uf3gt</i>	189.3	107.65

RCL7_23760_c0_g1	<i>MYB6</i>	664.67	354.33
RCL7_25096_c3_g1	<i>MYB11</i>	117.44	36.33
RCL7_27788_c0_g4	<i>bHLH1</i>	37.33	25.92
RCL7_20277_c0_g2	<i>bHLH2</i>	1.85	1.52
RCL7_22704_c0_g1	<i>WDR1</i>	19.48	11.32
RCL7_22679_c0_g1	<i>WDR2</i>	10.62	7.38

---

Table S3. Synonymous substitution rate ( $d_s$ ) along the branches leading to *Clarkia gracilis* and to its progenitor species, *C. amoena huntiana* and *C. lassenensis*. NA, not applicable.

Clade	Branch(es)		
	<i>C. gracilis</i>	<i>C. amoena huntiana</i>	<i>C. lassenensis</i>
<i>MYB6</i>	0.0371	NA	0.1415
<i>MYB1</i>	0.0218	0.0031	NA
<i>MYB11</i>	0.0248	0.0025	NA
<i>MYB12</i>	0.0085	0.0025	NA

Table S4. Voucher information for the *Clarkia* plants used in this study. RSABG, seeds obtained from Rancho Santa Ana Botanical Garden (Claremont, CA, USA). NFW, seeds from the collection of Norman F. Weeden. TRM, seeds provided by Talline R. Martins (University of Florida, Gainesville, FL, USA). LDG, seeds from the collection of Leslie D. Gottlieb.

Specimen	Voucher	Location	Phenotype
<i>C. amoena</i> <i>huntiana</i>	RSABG 15875	Mendocino County, CA	Central-spotted, pink-cupped
<i>C. amoena</i> <i>huntiana</i>	RSABG 15876	Marin County, CA	Central-spotted, pink-cupped
<i>C. lassenensis</i>	NFW 84b	Shasta County, CA	Basal-spotted, white-banded
<i>C. lassenensis</i>	NFW 152	Shasta County, CA	Basal-spotted, white-banded
<i>C. gracilis</i> <i>sonomensis</i>	TRM 14.6 (LDG 8513)	Sonoma County, CA	Central-spotted, pink-cupped
<i>C. gracilis</i> <i>sonomensis</i>	TRM 14.26 (LDG 8513)	Sonoma County, CA	Central-spotted, white-cupped
<i>C. gracilis</i> <i>albicaulis</i>	TRM 040 (Butte12)	Butte County, CA (Collected by B. Barringer)	Basal-spotted, white-banded

# A

```

CgsMYB1 ATGAATAAGGTAGGAGTTAGAAAGGTTGGTTGGACTGCAAAAGAAGATGC 50
CgaMYB1 ATGAATAAGGTAGGACTTAGAAAGGTTGGTTGGACTGCAAAAGAAGATGC 50
CahMYB1 ATGAATAAGGTAGGACTTAGAAAGGTTGGTTGGACTGCAAAAGAAGATGC 50
ClMYB1 -----GTAGTTGGACAAGAAAAGAAGATGC 25

CgsMYB1 CCTACTCAAGCAATGCGTTCAAACCTATGGAGAAGGCAACTGGCATCTAG 100
CgaMYB1 CCTACTCAAGCAATGCATTCAAACCTATGGAGAAGGCAATTGGCATCTAG 100
CahMYB1 CCTACTCAAGCAATGCGTTCAAACCTATGGAGAAGGCAACTGGCATCTAG 100
ClMYB1 CCTACTGAAGCAATGCGTTCAAACCTATGGAGAAGGTAAGTGGCATCTAG 75

CgsMYB1 TCCCCGATAGAGCAGGTTTGAATAGGTGCAGAAAAAGTTGCCGATTAAAG 150
CgaMYB1 TCCCCGATCGAGCAGGTTTGAATAGGTGCAGAAAAAGTTGCCGATTAAAG 150
CahMYB1 TCCCCGATAGAGCAGGTTTGAATAGGTGCAGAAAAAGTTGCCGATTAAAG 150
ClMYB1 TCCCCGATAGAGCAGGTTTGAATAGGTGCAGAAAAAGTTGCCAGATTGAGG 125

CgsMYB1 TGGCTTAACTATCTGAAACCAGGCTTAAACCGAGAAAGAGTTCCAAGAAGA 200
CgaMYB1 TGGCTTAACTATCTGAAACCAGGCTTAAACCGAGAAAGAGTTCCAAGAAGA 200
CahMYB1 TGGCTTAACTATCTGAAACCAGGCTTAAACCGAGAAAGAGTTCCAAGAAGA 200
ClMYB1 TGGCTTAACTATCTAAAGCCAGGCTATAAATCGAGGAGAGTTCCAAGAAGA 175

CgsMYB1 TGA AATTGACTTGATTATTAGGCTTCACAAGCTTTTGGCAATAAATGGT 250
CgaMYB1 TGA AATCGACTTGATTATTAGGCTTCACAAGCTTTTGGCAATAAATGGT 250
CahMYB1 TGA AATTGACTTGATTATTAGGCTTCACAAGCTTTTGGCAATAAATGGT 250
ClMYB1 TGA AATTGACTTGATTATTAGGCTTCACAAGCTTTTGGCAATAAATGGT 225

CgsMYB1 CACTCATAGCAGGAAGACTTCCTGGAAGAACAAGCAATGATATAAAGAAT 300
CgaMYB1 CACTCATAGCAGGAAGACTTCCTGGAAGAACAACAATGATATAAATAAT 300
CahMYB1 CACTCATAGCAGGAAGACTTCCTGGAAGAACAAGCAATGATATAAAGAAT 300
ClMYB1 CACTCATAGCAGGAAGACTTCCTGGAAGAACAAGCAATGATATAAAGAAT 275

CgsMYB1 TACTGGTAGACCCATATTGCCAAGAAATTA-----CCGCAAGAACC 341
CgaMYB1 TACTGGTAGACCCATATTGCCAAGAAATTA-----CCGCAAGAACC 341
CahMYB1 TACTGGTAGACCCATATTGCCAAGAAATTA-----CATGCAAGAACC 341
ClMYB1 TACTGGAAATAGTCATATTGTCAAGAAATTAGCCAAAGCTGCTGCACAACC 325

CgsMYB1 AGTTTATATCGCAGGAAAATTCAGTCAAAACACACGCCATAATAAGGCCTA 391
CgaMYB1 AATTATATCGCAGGAAAATTCAGTCAAAACACACGCCATAATAAGGCCTA 391
CahMYB1 AATTATATCGCAGGAAAATTCAGTCAAAACACACGCCATAATAAGGCCTA 391
ClMYB1 AATTATATTCGCAGGAAAATTCAGTCAAAAGAACACGTCATAATAAAGCCTA 375

CgsMYB1 TTGCAGGAAGACCAGACCAAAGGGATGCAGTTTGGTATGAA----- 431
CgaMYB1 TTGCAGGAAGACCCTACCAAAGGGATGCAGTTTGGTATGAA----- 431
CahMYB1 TTGCAGGAAGACCAGACCAAAGGGATGCAGTTTGGTATGAA----- 431
ClMYB1 TTGCAGGAAGACCTTACCAAAGGGATGAAAGTTTGGCATGAAATATCGTCTTA 425

CgsMYB1 --TGTACCTCTGATCAGCCTCCACCACCTCTGGAGAATTGGTCGGACTG 479
CgaMYB1 --TGTACCTCTGATCAGCCTCCACCACCTCTGGAGAATTGGTCGGACTG 479
CahMYB1 --TGTACCTCTGATCAGCCTCCACCACCTCTGGAGAATTGGTCGGACTG 479
ClMYB1 GGTGTCACCTCTGATCAGCCTCCACCACCTTGGGAGAATTGGTGGGACAG 475

CgsMYB1 GTTGATGATGGACGATGATAATAATTATTATGATAATGGTGGTGGTTGTT 529
CgaMYB1 GTTGATGATGGACGATGATAATAATTATTATGACAAATGGTGGTGGTTGTT 529
CahMYB1 GTTGATGATGGACGATGATAATAATTATTATGATAATGGTGGTGGTTGTT 529
ClMYB1 GTTGATGATGGGTGATGATAATAATTATTATGATTTTGGTGGTGGTTGTT 525

```

Fig. S1. (to be continued on the next page)



```

CgsMYB1  CTGCCTCGGAAGGCCACTGCACCACCGCTGCCAATGGCTGCTACGACATG 579
CgaMYB1  CCGCCTCGGAAGGCCACTGCACCACCGCTGCCAATGGCTGCTACGACATG 579
CahMYB1  CTGCCTCGGAAGGCCACTGCACCACCGCTGCCAATGGCTGCTACGACATG 579
ClMYB1   CCGCCTCAGGAAGGCCACTGCACCACCGCTACCAATG---GCTACCACCTG 572

CgsMYB1  CAGATAGAAATCTC---CGTGGCGGCAATAGTAGGCGGATCACCGAGGG 626
CgaMYB1  CAGATAGAAATCTC---CGTGGCGGCAATAGGAGGCGGATCACCGAGGG 626
CahMYB1  CAGATAGAAATCTC---CGTGGCGGCAATAGTAGGCGGATCACAGAGGG 626
ClMYB1   CAGATAGGAATCTCCGTCGTGTGCGGCAGAGGA-----ACTGAGAG 613

CgsMYB1  TGGGAAATAACCCGGATGAGCTTTATTTTGATGACATATTTTAA 669
CgaMYB1  TGGGAAATAACCCGGATGAGCTTTATTTTGATGACATATTTTAA 669
CahMYB1  TGGGAAATAACCCGGATGAGCTTTATTTTGATGACATATTTTAA 669
ClMYB1   TGGGTAATAACCC----- 625

```

Fig. S1. (to be continued on the next page)

**B**

```

CgsMYB6 ATGGGTGGTGTTCCTTGGACTGAAGAAGAGGATCTTCTGCTTAAGAAATG 50
CgaMYB6 ATGGGTGGTGTTCCTTGGACTGAAGAAGAGGATCTTCTGCTTAAGAAATG 50
CahMYB6 ATGGGTGGTGTTCCTTGGACTGAAGAAGAGGATCTTCTGCTTAAGAAATG 50
ClMYB6 -----CTGAAGAAGAGGATCTTCTGCTTAAGAAATG 31

CgsMYB6 CGTCGAGCAGTTCGGCGAAGGGAAATGGCACCGGGTTCCGCTTTTAGCCG 100
CgaMYB6 CGTCGAGCAGTTCGGCGAAGGGAAATGGCACCGGGTTCCGCTTTTAGCCG 100
CahMYB6 CGTCGAGCAGTTCGGCGAAGGGAAATGGCACCGGGTTCCGCTTTTAGCCG 100
ClMYB6 CGTCGAGCAGTTCGGTGAAGGGAAATGGCACCGGGTTCCGCTTTTAGCCG 81

CgsMYB6 GTCTGAACAGGTGCAGGAAGAGTTGCAGACTGAGGTGGCTGAATTATCTT 150
CgaMYB6 GTCTAAACAGGTGCAGGAAGAGTTGCAGACTGAGGTGGCTGAATTATCTT 150
CahMYB6 GTCTGAACAGGTGCAGGAAGAGTTGCAGACTGAGGTGGCTGAATTATCTT 150
ClMYB6 GTCTAAACAGGTGCAGGAAGAGTTGCAGACTGAGGTGGCTGAACCTACTC 131

CgsMYB6 CGACCGAATATCAAGAGAGGGAGCTTACTCAAGATGAAGTCGAGCTCAT 200
CgaMYB6 CGACCGAATATCAAGAGAGGGAGCTTACTCAAGATGAAGTCGAGCTCAT 200
CahMYB6 AGACCGAATATCAAGAGAGGGAGCTTACTCAAGATGAAGTCGAGCTCAT 200
ClMYB6 CGACCGAATATCAAGAGAGGGAGCTTACTCAAGATGAAGTCGAGCTCAT 181

CgsMYB6 CATCAAGCTCCATTAAGCTTGTGCGGAATCGGTGGTCGATGATTGCGGAA 250
CgaMYB6 CATCAAGCTCCATTAAGCTTGTGCGGAATCGGTGGTCGATGATTGCGGAA 250
CahMYB6 CATCAAGCTCCATTAAGCTTGTGCGGAATCGGTGGTCGATGATTGCGGAA 250
ClMYB6 CATCAAGCTCCACAAGCTTGTGCGGAATCGGTGGTCGATGATTGCTGGAA 231

CgsMYB6 GACTCCC GGGAAGAACAGCTAACGATGTCAAGAACTTTTGGAACTGTTCAT 300
CgaMYB6 GACTCCCTGGAAGAACAGCTAATGATGTCAAGAACTTCTGGAACTGTTCAT 300
CahMYB6 GACTCCC GGGAAGAACAGCTAACGATGTCAAGAACTTTTGGAACTGTTCAT 300
ClMYB6 GACTCCC GGGAAGAACAGCGAACGATGTCAAGAACTTTTGGAACTGCCAT 281

CgsMYB6 CTAAGCAAAAAGCTGACTGCCGAACAAATGAGCATCGACCCTGAACAGAG 350
CgaMYB6 CTAAGCAAAAAGCTGACTGCCGAACAAATGAGCATCGACCCTGAACGAAC 350
CahMYB6 CTAAGCAAAAAGCTGACTGCCGAACAAATGAGCATCGACCCTGAACAGAG 350
ClMYB6 CTGAGCAAAAAGCTGACTGCCGAACAAATGAGCATCGACCCTGAACGGAG 331

CgsMYB6 AATAGACAACCTAGTCCCAGTATAATAATGCCCAACC GCGAAA CCGTACAT 400
CgaMYB6 CATAGATAGCCTAGTCCCAGTATAATAATGCCCAACC GCGAAA CCGTGGAT 400
CahMYB6 AATAGACAACCTAGTCCCAGTATAATAATGCCCAACC GCGAAA CCGTACAT 400
ClMYB6 CATAGACAAGTCTAGTCCCAGTATAATAATGCCCAACC GGTAAA ATCGGTAT 381

CgsMYB6 CCGTCTCAGAGAAAACCGAACAAAAGAGACCAAGAAGTTGGGACTTCTATA 450
CgaMYB6 CCATCTCAGAGAAAACCGAACAAAAGAGACCAAGAAGTTGGTACCTCTATA 450
CahMYB6 CCGTCTCAGAGAAAACCGAACAAAAGAGACCAAGAAGTTGGGACTTCTATA 450
ClMYB6 CCGTCTCAGAGAAAACCGAACAAAAGAGACCAAGAAGTTGGGACTTCTATA 431

CgsMYB6 GTGACACTTCCAAGTGTTCGGTGAAGATGGAGCAATGAA GCGAGTTCAAGT 500
CgaMYB6 GTGACACTTCCAAGTGTTCGGTGAAGATGGAGCAATGAA GCGAGTTCAAGT 500
CahMYB6 GTGACACTTCCAAGTGTTCGGTGAAGATGGAGCAATGAA GCGAGTTCAAGT 500
ClMYB6 GTGACACTTCCAAGCGTTCGGTGAAGATGGAGCAATGAA GCAATTCAAGT 481

CgsMYB6 TATCGATGATGGAAAGACGAACCCACCGAATCAAGAACACAGGTGTCAGTTC 550
CgaMYB6 TATCGATGATGGAAAGACGAACCCACCGAATCAAGAACACAGGTGTCAGTTC 550
CahMYB6 TATCGATGATGGAAAGACGAACCCACCGAATCAAGAACACAGGTGTCAGTTC 550
ClMYB6 TATCGATCATGGAAAGATGAACCCGTCGAATCAAGAACACAGGTGTCAGTTC 531

```

Fig. S1. (to be continued on the next page)

```

CgsMYB6 TTGGAGATCTTCCGGTGAATTCCAGTTCGACGAATGTAGATTAGACGGG 600
CgaMYB6 TTGGAGATCTTCCAGGTGAATTCCAGTTCGATGAATGTAGATTAGACGGG 600
CahMYB6 TTGGAGATCTTGCAGGTGAATTCCAGTTCGATGAATGTAGATTAGACGGG 600
ClMYB6 TTGGAGATATTTCGGGAGAATTCCAGTTCGACGAATGTAGATTAGACGGG 581

CgsMYB6 ATTAGCAGCAGCAACAGCAGGAAATGGGACTGGGACGATCTGCTTATGGA 650
CgaMYB6 ATTAGCTGCAGCAACAGCAGGAAATGGGACTGGGACGATTTGCTCATGGA 650
CahMYB6 ATTAGCAGCAGCAACAGCAGGAAATGGGACTGGGACGATCTGCTCATGGA 650
ClMYB6 ATAAGCAGCAGCAACAGCAGAAAATGGGATGGGATGATCTACTCATGGA 631

CgsMYB6 TATGGATATCGATCTGTGGAATAACTCTCTGTAA 684
CgaMYB6 TATGGATATCGATCTGTGGAATAACTCTCTGTAA 684
CahMYB6 TATGGATATCGATCTGTGGAATAACTCTCTGTAA 684
ClMYB6 CATGGATAT----- 640

```

Fig. S1. (to be continued on the next page)

C

```
CgsMYB11 ATGAAGGGAGAA TTAAGGAAG GGTGTT TGG AATGCAGAAGAAGATGCTCT 50
CgaMYB11 ATGAAGGGAGAA TTAAGGAAG GGTGCT TGG AATGCAGAAGAAGATGCTCT 50
CahMYB11 ATGAAGGGAGAA TTAAGGAAG GGTGCT TGG AATGCAGAAGAAGATGCTCT 50
ClMYB11 ----- GGTG CATGG AATGCAGAAGAAGATGCTCT 29

CgsMYB11 CCTCAAGCAATGCATTCAAAC TTATGGAGAAGG AAAATGG CATCTTGTTTC 100
CgaMYB11 CCTCAAGCAATGCATTCAAAC TTATGGAGAAGG AAAATGG CATCTTGTTTC 100
CahMYB11 CCTCAAGCAATGCATTCAAAC TTATGGAGAAGG AAAATGG CATCTTGTTTC 100
ClMYB11 CCTCAAGCAATGCATTCAAAC TTATGGAGAAGG CAAATGG TATCTTGTTTC 79

CgsMYB11 CTGCA AGAACAGG ACTCA ATAGGTGCAGAAAAAGTTGCAG GTTGAGGTGG 150
CgaMYB11 CTACA AGAACAGG ACTCA ATAGGTGCAGAAAAAGTTGCAG GTTGAGGTGG 150
CahMYB11 CTGCA AGAACAGG ACTCA ATAGGTGCAGAAAAAGTTGCAG GTTGAGGTGG 150
ClMYB11 CCGCT AGAACAGG GCTGA ATAGGTGCAGAAAAAGTTGCAG ATTGAGGTGG 129

CgsMYB11 CTC AACTATCTG AAGCCGGGCATAAACCGTAG AGAGT TTTCAAGAAGATGA 200
CgaMYB11 CTC AACTATCTG AAGCCGGGCATAAACCGTAG AGAGC TTTCAAGAAGATGA 200
CahMYB11 CTC AACTATCTG AAGCCGGGCATAAACCGTAG AGAGC TTTCAAGAAGATGA 200
ClMYB11 CTC AACTATCTT AAACCGGGCATAAACCGTAA AGAGC TTTCAAGAAGATGA 179

CgsMYB11 AGTTGACTTG ATC ATCAGGCTTCATAAGCTTCTTGG CAATAGATGGTCAC 250
CgaMYB11 AGTTGACTTG GTC ATCAGGCTTCATAAGCTTCTTGG CAATAGATGGTCAC 250
CahMYB11 AGTTGACTTG ATC ATCAGGCTTCATAAGCTTCTTGG CAATAGATGGTCAC 250
ClMYB11 AGTTGACTTG ATC ATCAGGCTTCATAAGCTTCTTGG AAATAGATGGTCAC 229

CgsMYB11 TTATTGCA AGGAAG GCTACC GGAAGAACATCCACCCA AGTAAAGAATTAC 300
CgaMYB11 TTATTGCG GGAAG GCTTCC GGAAGAACATCCACCCA AGTAAAGAATTAC 300
CahMYB11 TTATTGCA AGGAAG GCTTCC GGAAGAACATCCACCCA AGTAAAGAATTAC 300
ClMYB11 TTATTGCA AGGAAG ACTTCC GGAAGAACATCCACCCA CGTAAAGAATTAC 279

CgsMYB11 TGG AATGCCCAT ATAGCTAAAAAGTGGAGATCGTCATC CAAAGCTGCACC 350
CgaMYB11 TGG AATGCCCGTACAGCTAAAAAGTGGAGATCGTCATC CAAATCTGCAGC 350
CahMYB11 TGG AATGCCCAT ATAGCTAAAAAGTGGAGATCGTCATC CAAAGCTGCAGC 350
ClMYB11 TGG AATGCCCAT ATAGCTAAAAAGTGGAGATCGTCATC AAA-----AGT 323

CgsMYB11 TGCAGAATCAAAAATCATCATC TTATAAAGGAA ---AACAGCAA GAATATTT 397
CgaMYB11 TGCAGAATCAAAAATCATCATC ---TAAAGGAA ---AACAAACA CAATATTT 394
CahMYB11 TGCAGAATCAAAAATCATCATC ---TAAAGGAA ---AACAGCAA GAATATTT 394
ClMYB11 TGCAGAATCAAAAATCATCATC ---AAAAGGAA AGCAGCAGCAA AAATATTT 370

CgsMYB11 CAGTAAACGTCATAA AGCCTATTGCTAG AAGAGCTCCCAAAATGATCGAT 447
CgaMYB11 CAGTAAACGTTATAA AGCCTATTGCTAG AAGAGCTCCCAAAATGATCGAT 444
CahMYB11 CAGTAAACGTCATAA AGCCTATTGCTAG AAGAGCTCCCAAAATGATCGAT 444
ClMYB11 CAGTCAACGTCATAA AGCCTATTGCTAG AAGAGCTCCCAAAATGATCAAG 420

CgsMYB11 TTTGGTATGACCATGAATAAACAATAATA TATGCAGCATGTCCACCTCTCA 497
CgaMYB11 TTTGGTATGATCATGAATAAACAATAATA TATGCAGCAGGTCCACCTCTCA 494
CahMYB11 TTTGGTATGATCATGAATAAACAATAATA TATGCAGCATGTCCACCTCTCA 494
ClMYB11 TTTGGT-----ATGAATGATAAATAATAA ATGCAGCAGGTCCACCTCTCA 464

CgsMYB11 GCTGCCGC-----CTCATCTTAGATTGTACCGGAATAATAAATACAG 538
CgaMYB11 GCTGCCGC-----CTCATCTTAGATTGTACCGAAATAATAAATACAG 535
CahMYB11 GCTGCCGC-----CTCATCTTAGATTGTACCGAAATAATAAATACAG 535
ClMYB11 GCGGCCGC CGCCTCCATCTCATCTTAGATTGTACGAAATAATAAATGAAG 514
```

Fig. S1. (to be continued on the next page)

```

CgsMYB11 ACAAGGGAGGTGATGAATTGGTTGCAGAGATTATTGGACGATGATGATTTA 588
CgaMYB11 ACAAGGAGGTGATGAATTGGTTGGAGAGATTATTGGACGATGATGATTTA 585
CahMYB11 ACAAGGAGGTGATGAATTGGTTGCAGAGATTATTGGACGATGATGATTTA 585
ClMYB11 ACAACGAGGTGATGAATTGGTTGGAGAGGTTATTGGACGATGATGATTTA 564

CgsMYB11 ATTGGT TTT GGAGGAGAC GGTGGCT CTGCCGCCTCAGAAGGCCACTGTAC 638
CgaMYB11 ATTGGT GGT GGAGGAGAC GGTGGCT CTGCCGCCTCAGAAGGCCACTGTAC 635
CahMYB11 ATTGGT GGT GGAGGAGAC GGTGGCT CTGCCGCCTCAGAAGGCCACTGTAC 635
ClMYB11 ATTGGT GGA GGAGGAGAC AGTGGCT GTGCCGCCTCAGAAGGCCACTGTAT 614

CgsMYB11 CACCCTAGGTGGTGGGTAATGTCGACTCCATGA 672
CgaMYB11 CACCCTAGGTGGTGGGTAATGTCGACTCCATGA 669
CahMYB11 CACCCTAGGTGGTGGGTAATGTCGACTCCATGA 669
ClMYB11 CGTCCTAGAAAGTGGGTAATATTCGACTGCATGA 648

```

Fig. S1. (to be continued on the next page)

**D**

```

CgsMYB12 ATGAAGGAAGGTCTAAGGAA GGGTGCTTGGAGTG CAGAAGAAGATGCTCT 50
CgsMYB12W ATGAAGGAAGGTCTAAGGAA GGGTGCTTGGAGTG CAGAAGAAGATGCTCT 50
CahMYB12 ----- GGGTGCTTGGAGTG CAGAAGAAGATGCTCT 30
ClMYB12 ATGAACGGAGGTCTAAGGAA GGGTGCTTGGAGTG AAGAAGAAGATGCTCT 50

CgsMYB12 CCTCAAGCAATGTATTCAAATTTATGGAGAAGGCAAATGGCATCTTGTTTC 100
CgsMYB12W CCTCAAGCAATGTATTCAAATTTATGGAGAAGGCAAATGGCATCTTGTTTC 100
CahMYB12 CCTCAAGCAATGTATTCAAATTTATGGAGAAGGCAAATGGCATATTGTTTC 80
ClMYB12 ACTCAGGCAATGCATTCAAACCTTATGGAGAAGGCAAATGGCATCTTGTTTC 100

CgsMYB12 CCGCCAGAGCAGGGCTAAATAGGTGTAGAAAAGGTTGCAGATTGAGGTGG 150
CgsMYB12W CCGCCAGAGCAGGGCTAAATAGGTGTAGAAAAGGTTGCAGATTGAGGTGG 150
CahMYB12 CCGCCAGAGCAGGGCTAAATAGGTGTAGAAAAGGTTGCAGATTGAGGTGG 130
ClMYB12 CCGCTAGATCAGGGCTGAATAGGTGCAGAAAAGGTTGCAGGTTGAGGTGG 150

CgsMYB12 CTCAACTATCTGAAGCCAGGCATAAACCTAAAGAGCTTCAAGATGATGA 200
CgsMYB12W CTCAACTATCTGAAGCCAGGCATAAACCTAAAGAGCTTCAAGATGATGA 200
CahMYB12 CTCAACTATCTGAAGCCAGGCATAAACCTAAAGAGCTTCAAGATGATGA 180
ClMYB12 CTCAACTATTGAAGCCCGGCATAAACCTAACAGAGCTTCAAGATGATGA 200

CgsMYB12 AGTTGACTTGATCTC TCAAACCTCACAAGCTTCTTGGCAACAAATGGTCAC 250
CgsMYB12W AGTTGACTTGATCTC TCAAACCTCACAAGCTTCTTGGCAACAAATGGTCAC 250
CahMYB12 AGTTGACTTGATCTC TCAAACCTCACAAGCTTCTTGGCAACAAATGGTCAC 230
ClMYB12 AGTTGACTTGATCTC TCAAACCTCACAAGCTTCTTGGCAACAAATGGTCAC 250

CgsMYB12 TTATAGCAGGAAGACTTCCGGGAAGAACATGCAATTATATAAAGAATTAC 300
CgsMYB12W TTATAGCAGGAAGACTTCCGGGAAGAACATGCAATTATATAAAGAATTAC 300
CahMYB12 TTATAGCAGGAAGACTTCCAGGAAGAACATGCAATTATATAAAGAATTAC 280
ClMYB12 TTATTGCAGGAAGACTTCCCGGAAGAACATGCAACTACATAAAGAATTAC 300

CgsMYB12 TGGAACTCCAAATATTGCTGCTAAAAAGTGGAAATCAAGAGAAAAGCAGCA 350
CgsMYB12W TGGAACTCCAAATATTGCTGCTAAAAAGTGGAAATCAAGAGAAAAG-AGCA 349
CahMYB12 TGGAACTCCAAATATTGCTGCTAAAAAGTGGAAATCAAGAGAAAAGCAGCA 330
ClMYB12 TGGAACTCCCATTTTACTGCTAAAAAGTCTAAATCAAGAAAAGCTGAA 350

CgsMYB12 AGATTCGGTCAAAGTTAAAGCCATAAGGCCTATTGTACGAAGAGCTCCCA 400
CgsMYB12W AGATTCGGTCAAAGTTAAAGCCATAAGGCCTATTGTACGAAGAGCTCCCA 399
CahMYB12 AGATTCGGTCAAAGTCAAAGCCATAAGGCCTATTGTACGAAGAGCTCCCA 380
ClMYB12 CGATTCGGTCAAAGT-----CATAAGGCCTATTGTACGAAGAGCTCCATA 394

CgsMYB12 AAATAATCAACTTTGGAAATGAATAATAATAAT-----ATA--GGAACC 441
CgsMYB12W AAATGATCAACTTTGGAAATGAATAATAATAAT-----ATA--GGAACC 440
CahMYB12 AAATGATCAACTTTGGAAATGAATAATAATAATAATAACAATAATGGAAGC 430
ClMYB12 AAATGATCAAGTTCGGTAATAATAATAATAATAATAAC-ATA--GGGACC 441

CgsMYB12 ACCTCTCA-----GCCTCATTGTAACGG 464
CgsMYB12W ACCTCTCA-----GCCTCATTGTAACGG 463
CahMYB12 ACCTCTCA-----GCCTCATTGTAACGG 453
ClMYB12 ACCTCTCGGCCGCCGCTCACTGCCGCCAACACCTCCGCCTCATTGTAACGG 491

CgsMYB12 AATAACTAGGACGACGAGGTGATGAATTGGTTGGATAGGTTACTAATGG 514
CgsMYB12W AATAACTTGGGACGACGAGGTGATGAATTGGTTGGATAGGTTACTAATGG 513
CahMYB12 AATAACTACAGACGACGAGGTGATGAATTGGTTGGATAGGTTACTAATGG 503
ClMYB12 AATAAAATACAGACAACGAGGTGGTGAATTGGTTGGATAGGTTACTAGTGG 541

```

```

CgsMYB12  AC GATGATGATGATGTTTATGCTTTTGTAAAGGAGACGGGGCTGTACC 564
CgsMYB12W AC GATGATGATGATGTTTATGCTTTTGTAAAGGAGACGGGGCTGTACC 563
CahMYB12  AC --GATGATGATGTTTATGCTTTTGTAAAGGAGACGGTGGCTGTACC 550
ClMYB12   AC ----GATGATGTTTATGCTTTTGTGGAGGAGACAGTGGTGTGTCC 585

CgsMYB12  GCCTCACAAAGGCCACTGCTCCAACCGCAGGTGGTGGC GGGTGTATCGACGG 614
CgsMYB12W GCCTCACAAAGGCCACTGCTCCAACCGCAGTTGGTGGC GGGTGTATCGACGG 613
CahMYB12  GCCTCACAAAGGCCACTGCTCCAACCGCAGGTGGT---GGTGTATCGACAG 597
ClMYB12   GCCTCACAAAGGCCACTGTGCCATCGCAGGAGGT---GGTATATCGACGA 632

CgsMYB12  CAGTGTCTAGACGAGCTGTATATTGATCACGACATATTTTCAGTTGTAG 663
CgsMYB12W CAGTGTCTAGACGAGCTGTATATTGATCACGACATATTTTCAGTTGTAG 662
CahMYB12  CAATGTCTAGACGAGTTTATATT----- 622
ClMYB12   CAGTGTCTAGACATGCTTATATTGATCATCACATATTTTCAGCTGTAG 681

```

Fig. S1. Alignment of *Clarkia R2R3-MYB* sequences (A) *MYB1*, (B) *MYB6*, (C) *MYB11* and (D) *MYB12* from *C. gracilis sonomensis*, *C. g. albicaulis*, *C. amoena huntiana* and *C. lassenensis*. If the nucleotides at each column are not identical, this column is highlighted in a black background. The sequences retrieved from GenBank are: *CgsMYB1* (KX592432), *CgaMYB1* (KX592431), *ClMYB1* (KX592428), *CgsMYB6* (MT425534), *CgsMYB11* (MT425536), *CgsMYB12* (MT425538) and *CgsMYB12<sup>W</sup>* (MT425537). Other sequences were generated in this study and were deposited in GenBank under the accession numbers MT796894-MT796902. *CgsMYB12<sup>W</sup>* is the nonfunctional copy of *CgsMYB12*, with the presence of a deletion (highlighted in a gray background) that creates a premature stop codon (highlighted in a red background). The conserved MYB DNA-binding domain, consisting of R2 and R3 repeats, is indicated in pink and blue letters. The fragment corresponding to the conserved amino acid motif specific to the anthocyanin-regulating *R2R3MYBs* is indicated in orange letters.

**A** *C. amoena huntiana*

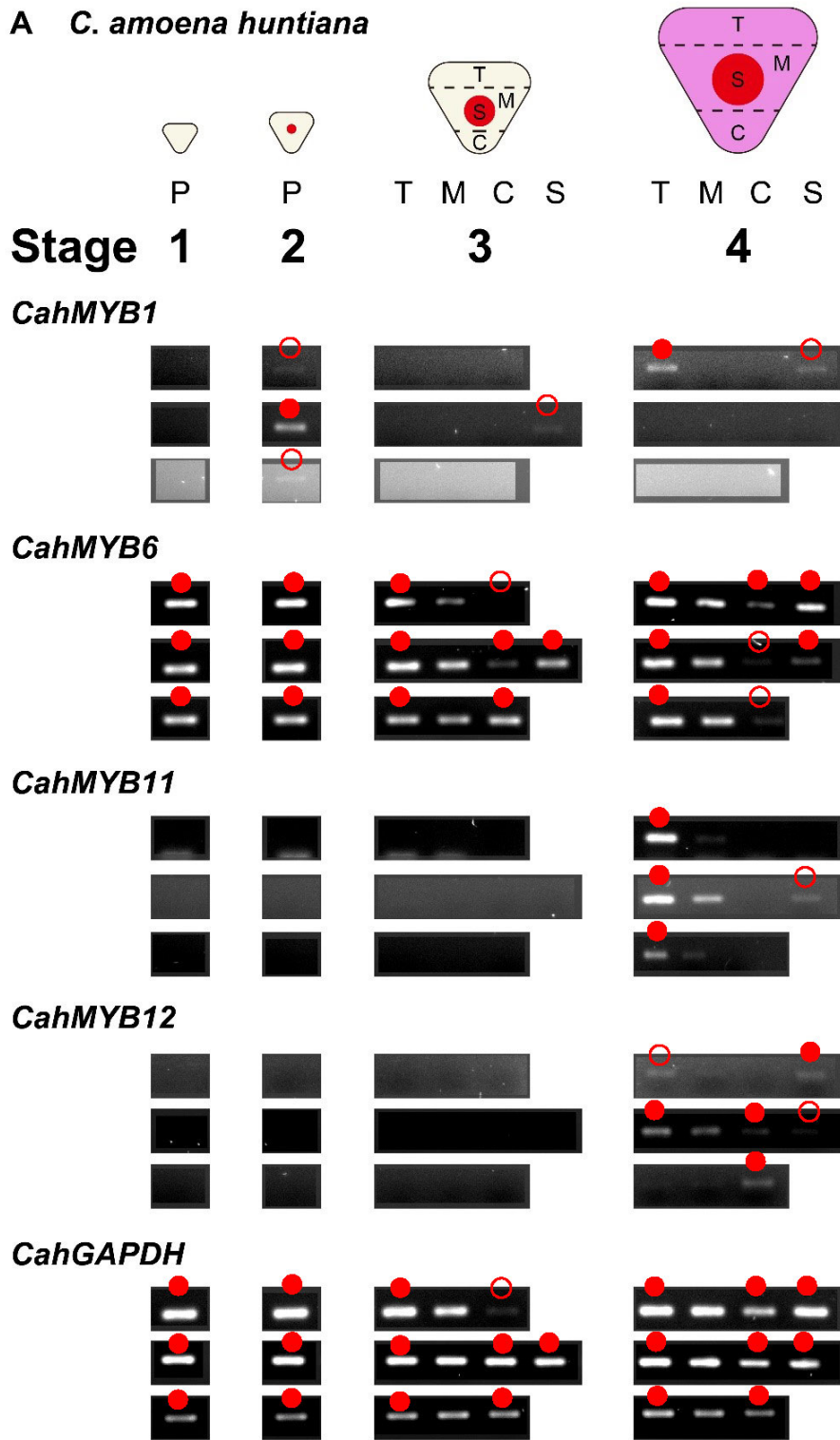


Fig. S2. (to be continued on the next page)



**B** *C. lassenensis*

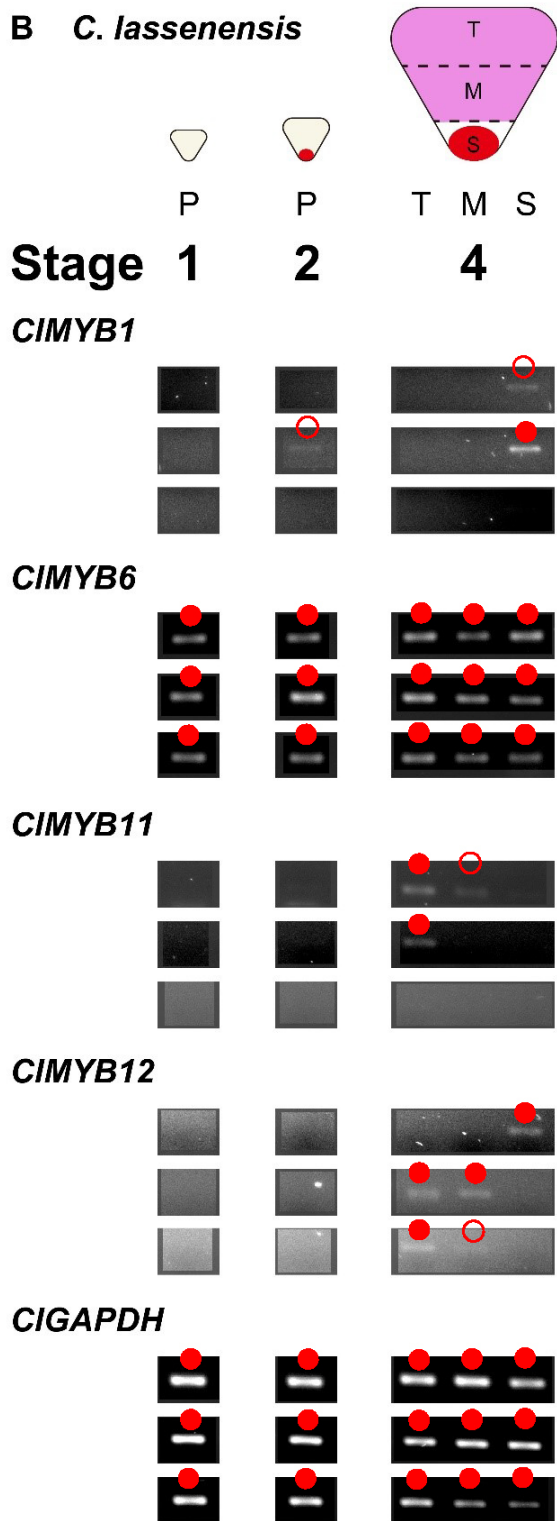


Fig. S2. (to be continued on the next page)

**C Pink-cupped *C. gracilis sonomensis***

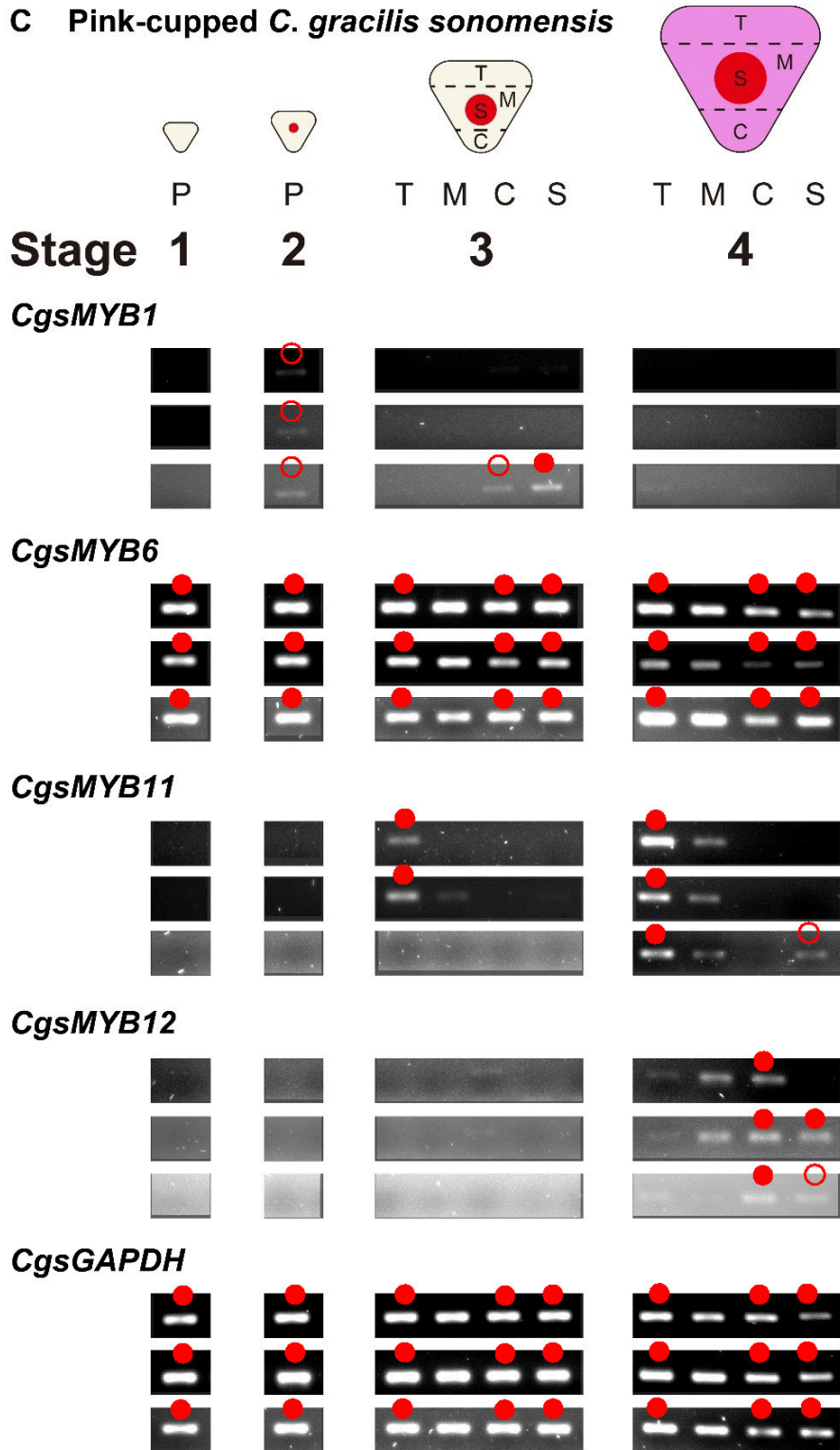


Fig. S2. (to be continued on the next page)

**D White-cupped *C. gracilis sonomensis***

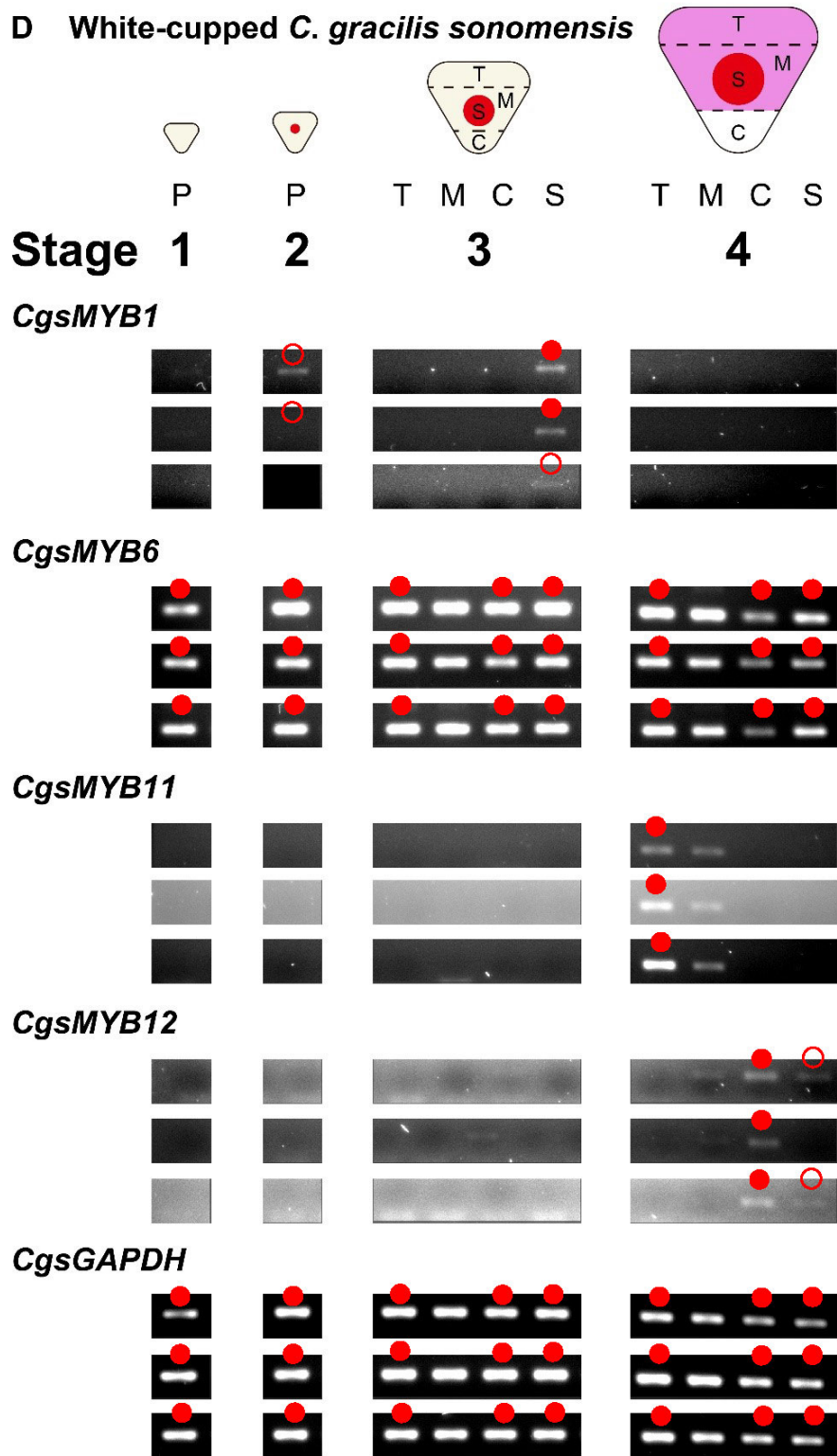


Fig. S2. (to be continued on the next page)

**E** *C. gracilis albicaulis*

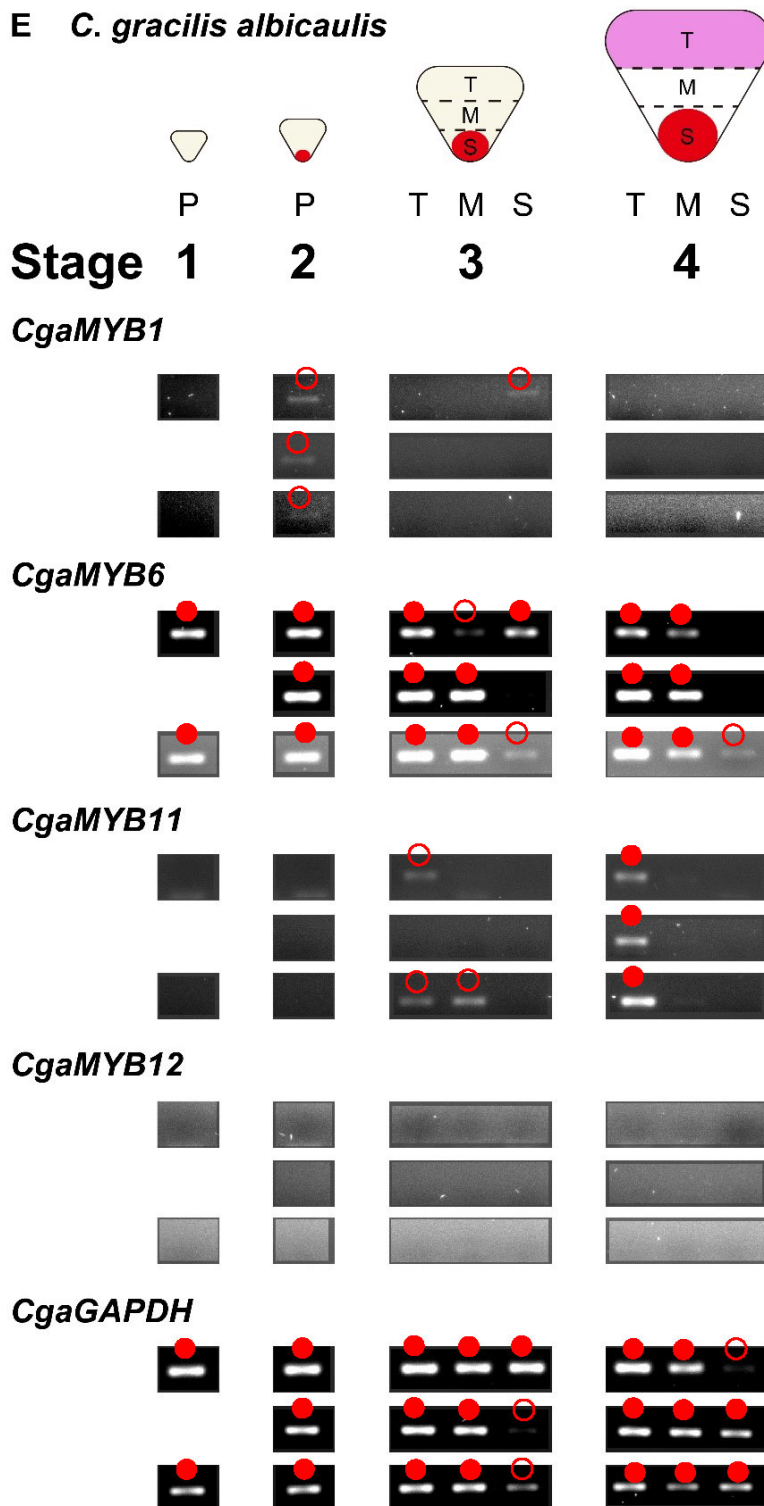


Fig. S2. Spatiotemporal expression of *MYB1*, *MYB6*, *MYB11* and *MYB12* in the petals of (A) *Clarkia amoena huntiana*, (B) *C. lassenensis*, (C) pink-cupped *C. gracilis sonomensis*, (D)

white-cupped *C. g. sonomensis*, (*E*) *C. g. albicaulis*. Flower buds were collected from three plants from each (sub)species/phenotypes. For Stages 1 and 2, the whole petals (P) were used. For Stages 3 and 4, the petals were dissected into sections according to coloration: T (top), M (middle, excluding the spot), C (cup) and S (spot). For *C. lassenensis*, Stages 2 and 3 were combined into Stage 2 due to a small petal size. A constitutively expressed gene *GAPDH* was included for cDNA quality control. Red circles indicate detected bands. Filled circles correspond to “+” and open circles to “(+)” in fig. 4.

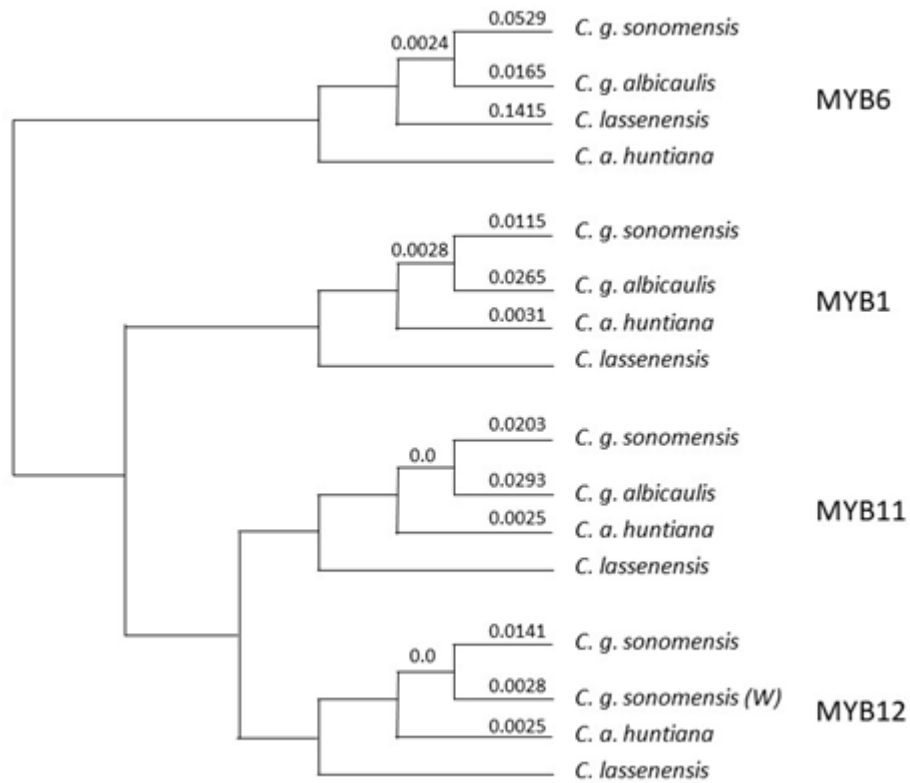


Fig. S3. Gene tree used to estimate  $d_S$  (synonymous substitution rate) along branches. Numbers above branches indicated estimated  $d_S$ . (W) indicates the nonfunctional allele of *MYB12* in *Clarkia gracilis sonomensis*.

## Methods S1. Transcriptomics

To perform a comprehensive survey on which transcripts were expressed differentially between the pink background (the most distal portion) and the white band (the central portion) of the *Clarkia gracilis albicaulis* petals, we sequenced RNA extracted from these two petal regions. The petal regions were collected by dissecting flower buds (approximately 1 day before flowering) from four plants. Total RNA of pink background and white band was individually extracted using Spectrum Plant Total RNA Kit (Sigma-Aldrich, St. Louis, MO, USA). RNA samples of the two petal regions for library construction were prepared by pooling equal amounts of RNA from each plant. Prior to library construction, RNA quality was examined using Bioanalyzer Agilent RNA 6000 Nano Kit (Agilent Technologies, Santa Clara, CA, USA). We followed the protocol described in Supporting Information Methods S2 in Lin and Rausher (2021) to construct, barcode and sequence the libraries.

Bioinformatic analyses, including transcriptome assembling, identification of candidate genes for anthocyanin production and estimation of gene expression, were conducted as described in Lin and Rausher (2021).

Raw sequence reads generated in this study were deposited under NCBI Bioproject PRJNA721169 (Sequence Read Archive accessions: SRR14226355 and SRR14226356). The assembled transcriptomes were deposited at the NCBI Transcriptome Shotgun Assembly under the accessions GJDS00000000 and GJDT00000000. FPKM data are available at <http://doi.org/10.5281/zenodo.4699974>.

### References cited in Supplementary Material

- Lin R-C, Rausher MD. 2021. *R2R3-MYB* genes control petal pigmentation patterning in *Clarkia gracilis* ssp. *sonomensis* (Onagraceae). *New Phytol.* 229(2):1147–1162.
- Martins TR, Jiang P, Rausher MD. 2017. How petals change their spots: *cis*-regulatory re-wiring in *Clarkia* (Onagraceae). *New Phytol.* 216(2):510–518.