

Isolation and characterisation of GtMYBP3 and GtMYBP4, orthologues of R2R3-MYB transcription factors that regulate early flavonoid biosynthesis, in gentian flowers

Takashi Nakatsuka, Misa Saito, Eri Yamada, Kohei Fujita, Yuko Kakizaki, Masahiro Nishihara

Gentian_MYBP3	1	MGRAPCCEKVGLKRGWRWTAEDEIITNYIQANGEGSWRSLPKNAGLLRCGKSCRLRWINYLRSDLKRGNISEQEEIIIIRLHANLGNRWS	90
Gentian_MYBP4	1	MTRAPCCEKVGLRGRWRWTAEDESLINYYIQHFGEGSWRLPKNAGLLRCGKSCRLRWINYLRSDIKRGNFTIIEEEETIIQLHRSLGNRWS	90
Arabisopsis_MYB12	1	VGRAPCCEKVGIKGRWRWTAEDEQILSNYIQSNGEGSWRSLPKNAGLLRCGKSCRLRWINYLRSDLKRGNIITPEEEELVVKLHSTLGNRWS	90
Arabidopsis_MYB11	1	MGRAPCCEKVGIKGRWRWTAEDERTISDYIQSNGEGSWRSLPKNAGLLRCGKSCRLRWINYLRSDIKRGNITPEEDIVVKLHSTLGTRWS	90
Arabidopsis_MYB11	1	MGRAPCCEKVGIKGRWRWTAEDEIITKYIQQTNGEGSWRSLPKNAGLLRCGKSCRLRWINYLRSDLKRGNIITSDEEIIIVKLHSLLGNRWS	90
Grape_MYBF1	1	MGRAPCCEKVGLKRGWRWTAEDEEVLYKVIQANGEGSWRSLPKNAGLLRCGKSCRLRWINYLRSDLKRGNISEEEEEEIIIKLHLASLGNRWS	90
Gerbera_MYB1	1	MGRAPCCEKVGLKRGWRWTAEDEIITNYIKTHGEGSWRSLPKNAGLLRCGKSCRLRWINYLRSDLKRGNIKSEDEDVITKLHASLGNRWS	90
Tomato_MYB12	1	MGRAPCCEKVGIKGRWRWTAEDEQILTNYITISNGEGSWRSLPKNAGLLRCGKSCRLRWINYLRSDLKRGNIITSQEEEDIKKLHATLGNRWS	90
Maize_P1	1	MGRAPCCEKVGLKRGWRWTAEDEQLIANYIAEHGEGSWRSLPKNAGLLRCGKSCRLRWINYLRADVKRGNISKEEEDIKKLHATLGNRWS	90
Maize_P2	1	MGRAPCCEKVGLKRGWRWTAEDEQLIANYIAEHGEGSWRSLPKNAGLLRCGKSCRLRWINYLRADVKRGNISKEEEDIKKLHATLGNRWS	90
Sorghum_Y1	1	MGRAPCCDKVGLKRGWRWTAEDEQILIANYIAEHGEGSWRSLPKNAGLLRCGKSCRLRWINYLRADVKRGNISKEEEDIKKLHATLGNRWS	90
R2 Repeat		R3 Repeat	
Gentian_MYBP3	91	VAGNLPGRTDNEIKNYWNNSHLRSRKIHT-----FRRPVCGRTADTAHALIMDVARAKLAMKKL	147
Gentian_MYBP4	91	MIAHLPGRTDNEVKNYWNNSHLRSKLYSFTTSSHNKTINGGGGSNLK---TTAVDLIKMSNNNTSSKRTKRHGGGGGRVRRSIATKYN	176
Arabisopsis_MYB12	91	LIAGHLPGRTDNEIKNYWNNSHLRSKLNHFIRKPISQDVSAVIMT-----NASSAPPPPQAKRRLGRTGRSAMPKFILNPKNHKTPN	173
Arabidopsis_MYB11	91	TIASNLPGRTDNEIKNYWNNSHLRSKLNHGFRKPTVANTV-----ENAPPPPKRPRGRTSRSAMPKFILNPKNHKTPN	163
Arabidopsis_MYB11	91	LIATHLPGRTDNEIKNYWNNSHLRSKITYAFTAVSGDGHNLVNDVLVLLKSCSSSSGAKNNNKTKKKKGRTTSRSSMKHHKQMVTAQSFCQ	180
Grape_MYBF1	91	MIAGQLPGRTDNEIKNYWNNSHLRSKXVHSFRLLTNEGPMV-----IDLAKVTAAHKRKVGRTSRWAAMKKNRSDKSIREDVNK	167
Gerbera_MYB1	91	LIAAHLPGRTDNEIKNYWNNSHLRSKIIPIPSRLLNTPATPVP-----AKLSSHPLNKRKGRTSRSAMSKINKTHKSSSRDIGA	166
Tomato_MYB12	91	LIAEHLSPGRTDNEIKNYWNNSHLRSKVDSSLRIPSDEKLPKA-----VVDLAKKGIPKPKISSISRPKNNCSNLLEKEALCCTN	168
Maize_P1	91	LIAHLPGRTDNEIKNYWNNSHLRSQIHTYRRKYTAGPDATIA-----IDMSKLQSAQRGGRTGRPPKIASARTKQADADQP	170
Maize_P2	91	LIAHLPGRTDNEIKNYWNNSHLRSQIHTYRRKYTAGPDATIA-----IDMSKLQSAQRGGRTGRPPKITSASRTKHSADADQP	170
Sorghum_Y1	91	LIAHFEPGRTDNEIKNYWNNSHLRSQIHTYRRKYTAAPDTVI-----IDMSKLHSAEKRRCGRTPGWSPKSSSANNTNTSSK	169
R3 Repeat		SG7 motif GRTxRSxMK	
Gentian_MYBP3	238	FESELWHDGIGEWELENSPA VSSNVNEGKLWNFENMQWEELAKETRALEEEE-----	288
Gentian_MYBP4	267	TTGITATTAAATSTPVEVGPDENEDECWFLLTTLFDQMQLSPLKLNNEGVMQDCGMANNDGNYDKGFSSLMSFQLTNEEST-----	348
Arabisopsis_MYB12	264	SEGARGFSDTWNQGNLDCLLQSCPVESFLNYDHQVNDASTDEFIDWDCVWQEGSDNNLWHEKENPDSMVSWL LDGDEATI-----	345
Arabidopsis_MYB11	254	SNNVKNSEPYGGMGSVGHKNIE TMADD FVDWVGLCNSEVFGVDEELLDW-----	324
Arabidopsis_MYB11	271	HDMTVEREREGFLSSNSNEENNDDKWWVGLCNSEVFGVDEELLDW-----	317
Grape_MYBF1	258	VSSGTEERASGATCSNKTTPFGGDVEGRNLSSNGDQTV DQWPSCSSPTSYFDDWNWESSAVNGQELWDEKEEMLSWLWEDSDGE-----	343
Gerbera_MYB1	257	NVTSEEERQNVSEPIISIIGGEEVINKGSCATSTSTDSCNVGDHNILSGPWDWKWNFDVEEGMLGLGV EDEDNLISWPWESTTT-----	342
Tomato_MYB12	259	ISSDDEKIKL LMDWQDNDELVWPTLPWELET DIVPSWPQWDDTDNLLQNCTNDNN-----	314
Maize_P1	261	RSAARPRWTTCSWTGMASRPICAGGRSRSTSTARSGRPPSRWKLLLLLRLRRPAPRTIASWRRSRLGSCP TRSDGSHRTDQTDQIIIG	350
Maize_P2	261	GEAQVDDLFDMDWDGF AAHLWGGEQDDHSAQLRQAAEP-----	299
Sorghum_Y1	260	ESEIDALMSMDAPLEGFD AVVGEAQAQVDDLFDMDMDWDGF AAHLWGGEQNDHIAELQQAAEPQATAAACTPNEHEPQVAAAAA-----	345
SG7-2 motif [W/x][L/x]LS			
Gentian_MYBP3	289	-----NALS WQWHTSGGGYGEENC DGTSGGEN DVKKQNAIVS [W/L]FS	329
Gentian_MYBP4	349	-----NNWDFNEAAFVYDGLWDDQDNS [L/W]V	376
Arabisopsis_MYB12	346	-----GNSNCNFGEPLDHDDES [L/V]A [W/L]LS	371
Arabidopsis_MYB11	325	-----FGEPLDIDEENKMAA [W/L]LS	343
Arabidopsis_MYB11	318	-----EFQGNVTCQSDDL [W/L]LS [D/G]EITLE	342
Grape_MYBF1	344	-----VECETFGGDLCEKQNAMVA [W/L]LS	367
Gerbera_MYB1	343	-----DSGN GDTDFVGE DLEKQNAMVA [W/L]LS	368
Tomato_MYB12	315	-----NYEEATTMEINNQNQHSTIVS [W/L]LS	338
Maize_P1	351	-----SRVLARSLPSRGSWFRWPNNWEKNSTARAVKPPCAPDVDACRVE [L/R]L	399
Maize_P2	300	-----MEAAA VAAAAAAAATAACTPDDRELEAFET [W/L]SOSF	335
Sorghum_Y1	346	-----CTPDEHEPQAAAAAAATCTPDEHGLEAFET [W/L]SOSF	383

Fig. S1. Alignment of P1-orthologue proteins in higher plants.

The position of the R2R3-DNA binding domain is indicated below the alignment.
The conserved SG7-1 and SG7-2 domains are highlighted by black boxes.

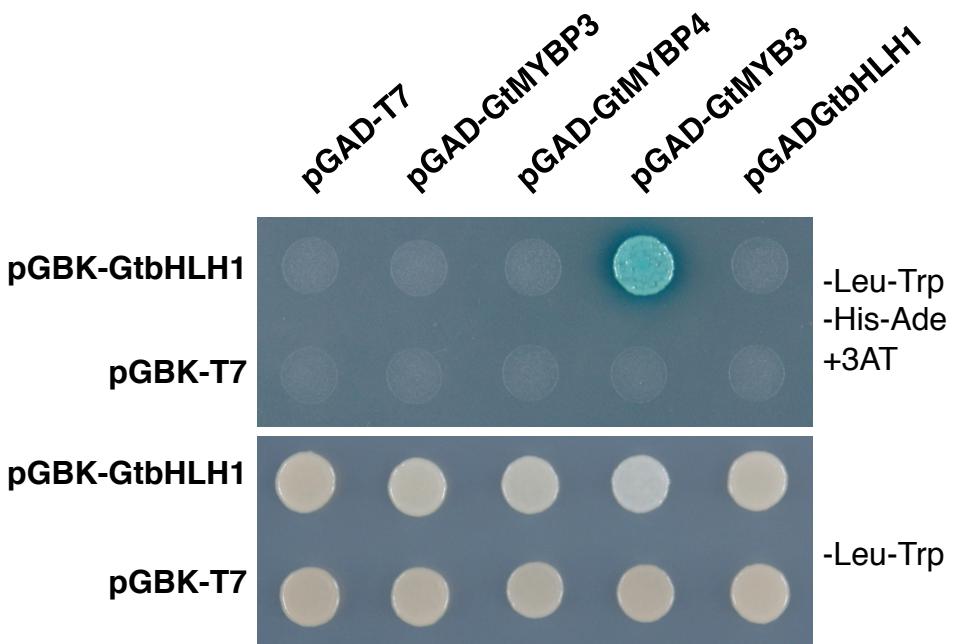


Fig. S2. Yeast two-hybrid analysis to examine the protein-protein interaction between GtMYBs and GtbHLH1.

The GtbHLH1 protein was fused to the GAL4 binding domain (BD) and assayed for its ability to bind the GtMYBs and GtbHLH1 fused to the GAL4 activation domain (AD). The transformed yeasts were grown on quadruple dropout medium (without leucine, tryptophan, histidine and adenine, upper) supplemented with 15 mM 3AT and on double dropout medium (without leucine and tryptophan, lower) at 30°C for 3 days. Protein-protein interactions are shown by yeast growth on quadruple dropout medium.

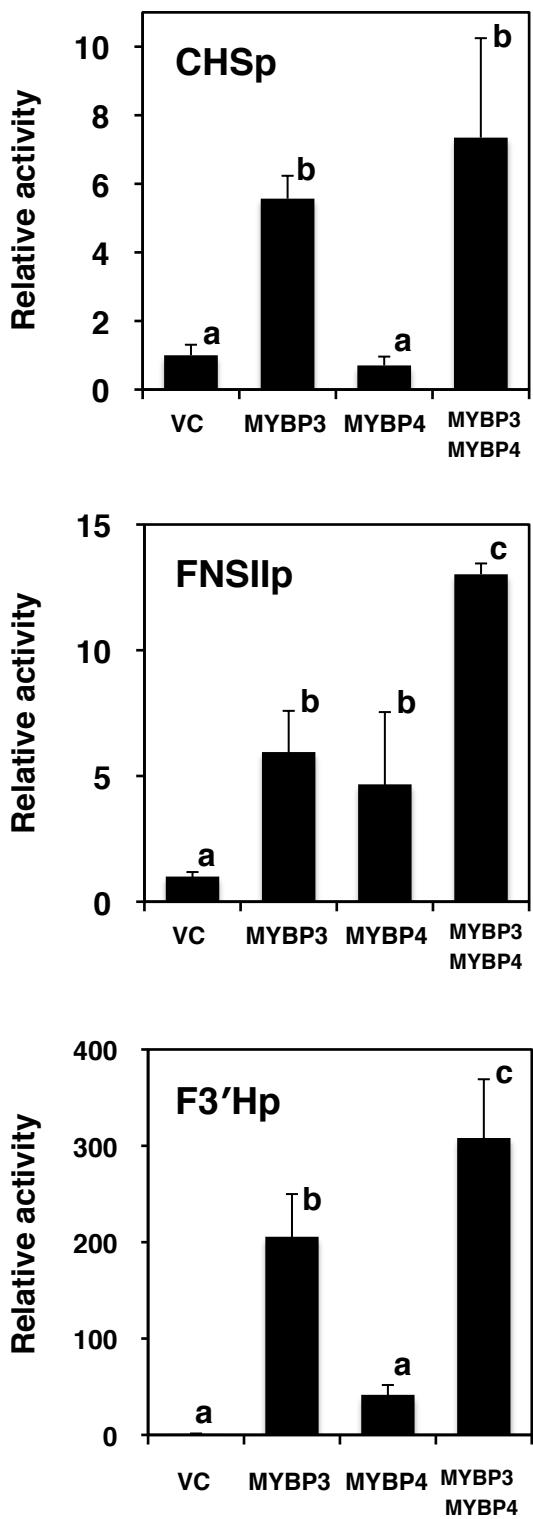


Fig. S3. Effect of co-transfection of *GtMYBP3* and *GtMYBP4* on promoter activities of *GtCHS*, *GtFNSII*, *GtF3'H*.

Transient expression assays were performed using protoplasts from *Arabidopsis* T87 cells, as described in Fig. 4B. Letters are the results from Tukey's multiple comparisons test where different letters represent a significant difference at $P<0.05$.

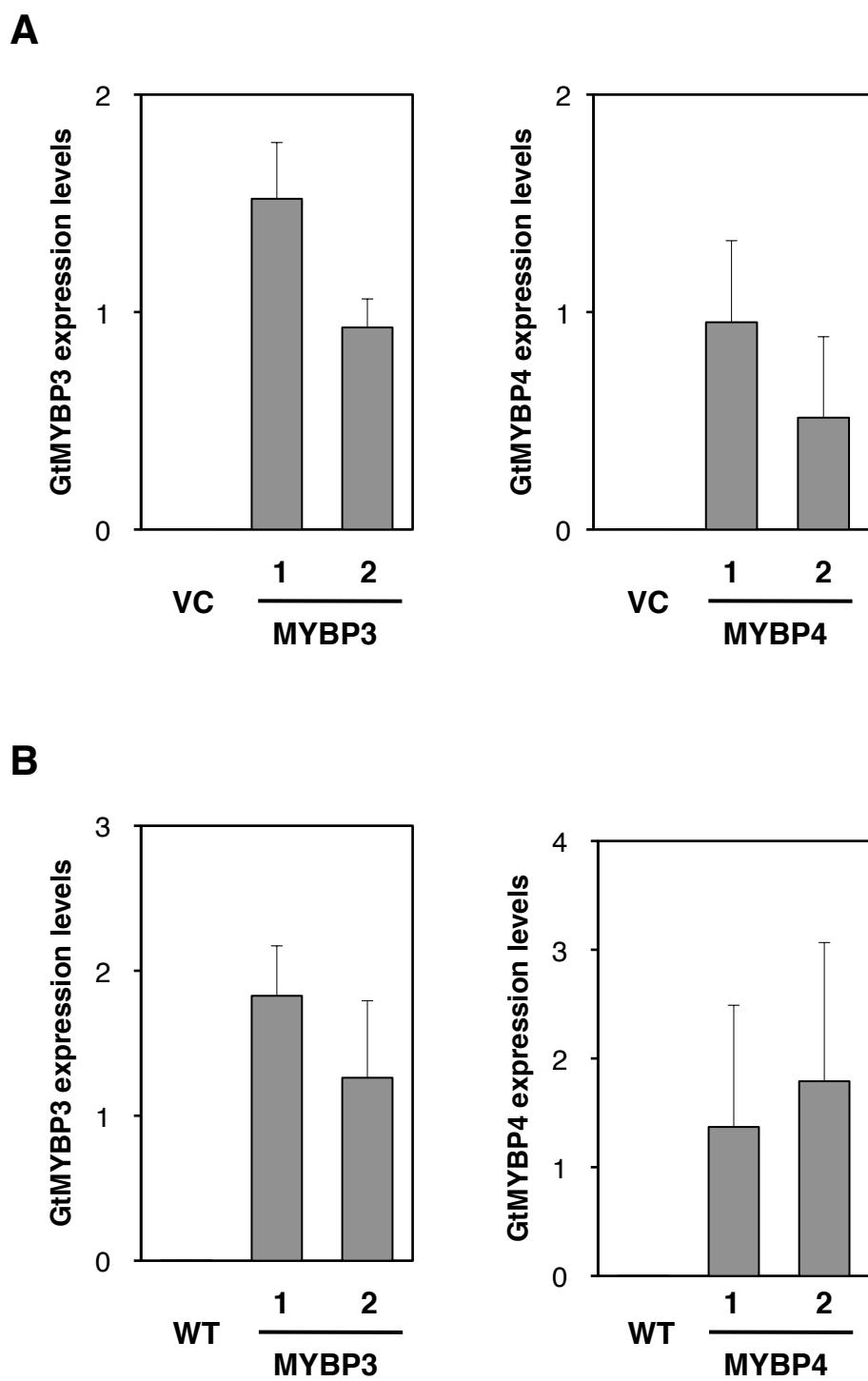


Fig. S4. Confirmation of the expression of transgenes in transgenic Arabidopsis and tobacco.

The expressions of *GtMYBP3* and *GtMYBP4* were investigated in transgenic 5-day-old Arabidopsis seedlings (A) and in transgenic tobacco petals (B). VC indicates vector control.

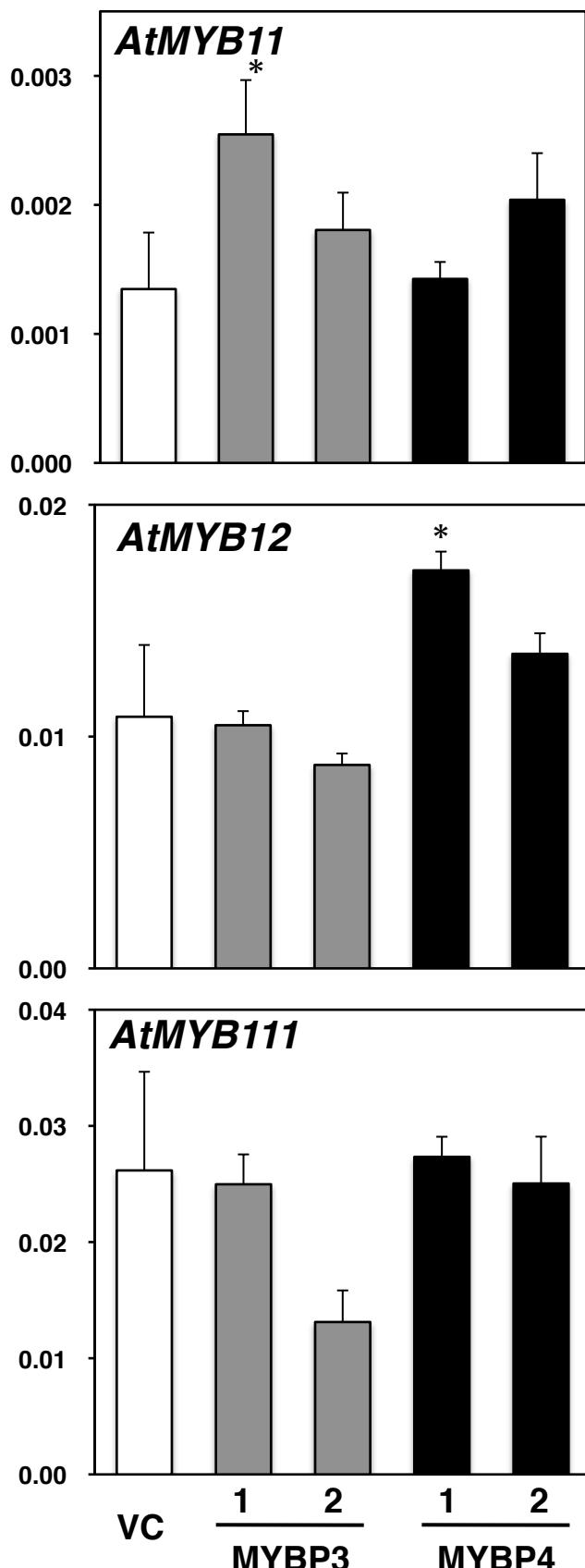


Fig. S5. Expression analyses of endogenous flavonol-specific transcription factor genes in transgenic *Arabidopsis*.

The effects of *GtMYBP3* and *GtMYBP4* overexpression on endogenous flavonol-specific R2R3MYB genes were investigated using qRT-PCR analyses in vector control and 5-day-old transgenic seedlings. The two independent transgenic lines shown in Fig. 6 were analysed. Asterisks (*, $P < 0.05$) represent statistically significant differences between the means for vector control and transgenic lines, as judged by Student's *t*-test.

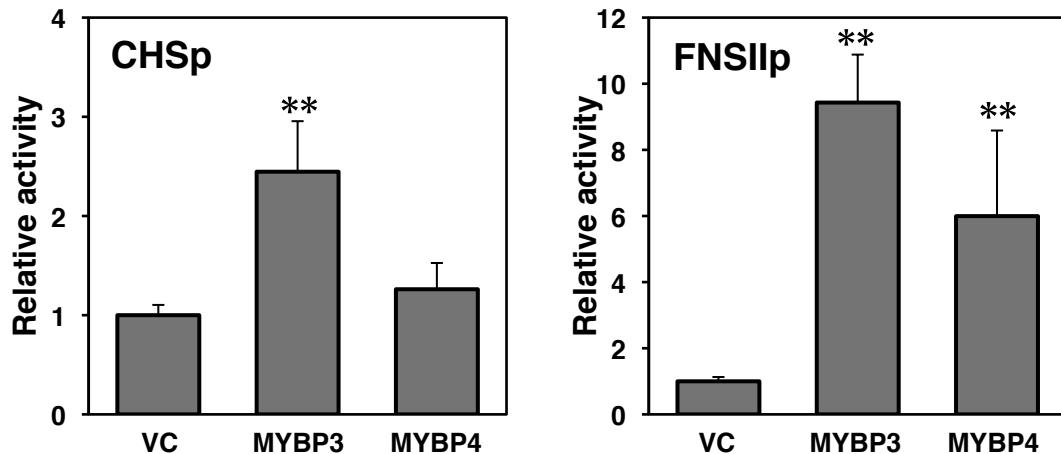


Fig. S6. Transient expression assay in gentian mesophyll protoplast.

The effect of GtMYBP3 and GtMYBP4 on promoter activities of two early flavonoid biosynthetic genes, *GtCHS* and *GtFNSII*, were investigated in mesophyll protoplasts of gentian. Transient expression assays were performed by the dual-Glo luciferase assay system, as described in Fig. 4B. Asterisks (**, $P < 0.01$) represent statistically significant differences between the means for negative control (pBI221) and tested genes, as judged by Student's *t*-test.

Table S1. Primers used in this study.

		Forward	Reverse
Degenerate PCR	Set 1	GRB TDM GRA ARG GTK CWT GGA	GCW ATH ARD GAC CAY CTR TT
	Set 2	AAR WSI TGY MGI YTI MGI TGG AYN AAY TA	CCA RTA RTT YTT IAM YTC RTT RTC
Probe	MYBP3	ATG CAA ATC TGG GTA ACA GGT GGT	CGT CTC TTT CGC TAA TTC TTC CCA
	MYBP4	TTC ATC CAA GCG GAC CAA GC	CCA TCC CCT CAT TAT TTA GC
Inverse PCR	FNSII pro	CGT TGT ATT GTA GTC TCC ACC CCT GAA CTA GC	GGA GGA TTG AGA GGA CGA AAA TGG AAG CAG AG
	F3'H pro	TCC TTA TAC GAG CAA TAG CGA GTG GCG GTC AT	TGA GGA AGA TGA AGA TGA AGA CGA CGA AAA GG

Table S2. Primers used for quantitative RT-PCR analyses

	<u>Primer sequence (5' - 3')</u>	
	Forward	Reverse
Arabidopsis		
<i>AtCHS</i>	CTAAGGATCTGCCGAGAACAA	CGGCTGTGATCTCAGAGCAG
<i>AtCHI</i>	GTCACCGGCCTCCTCCA	TGGATATCAAGGCCTCGGAC
<i>AtF3H</i>	TGAAGGAGCGTTGTCGTCA	TGAACCTCCCATTGCTAAAAA
<i>AtF3'H</i>	CCTCCACCTCCGACTAGGGT	TGCTCGGCCACGGATTAA
<i>AtFLS</i>	CAACATTCCGAGGTCCAACG	TCTTCGTCGGGATCGCTTAG
<i>AtDFR</i>	CGAGATGACGGCAGCTTG	AGCGGCGACATGGAAGAC
<i>AtMYB11</i>	TCGCCAATACCGTCGAGAAAT	CGGATCTGCTGGTTCTTCCA
<i>AtMYB12</i>	CGTAAAACGAAGAAAACGTCTGC	GCTTCTTATCAGCCCCAGCT
<i>AtMYB111</i>	CAATGTTCTCACAACCTAAGGAGC	CCAAAGACTCTCCTCAAAATTACCA
<i>AtACT2</i>	ACCCGATGGGCAAGTCATC	CGAGGGCTGGAACAAGACTTC
Tobacco		
<i>NtPAL</i>	TGCTTAACCACAATGTCACTCCA	CGAGATCACCAGAGGCGGT
<i>NtCHS</i>	GCCGGTGGCACGGTACT	ACTCGAGGCCCTTGTGT
<i>NtCHI</i>	TGAAGCAGTGTGGAATCCA	TTTCGGCGATACTACACTTGC
<i>NtF3H</i>	TTTTTACCCAAAGTGTCCACAGC	GGGTGATGGTTCTGGATCA
<i>NtF3'H</i>	TGGATTAACCCATTCAATTGGAT	TTCCAAAAGGCTAACACTTCTC
<i>NtDFR</i>	TGAGTTAAAGGCATCGATAAGGA	GAATTGAAACCCATATCCGTC
<i>NtANS</i>	TCCTCCACAATATGGTGCCTG	GGGTGTCCCCAATATGCATG
<i>NtFLS</i>	GGCCTAAAAATCCTCCCTCCT	TTCTCCACAACCTCTCGCAGC
<i>NtUBQ</i>	AAGATTCAAGGACAAGGAAGGCA	AGCTGCTTACCTGCGAAAATCA