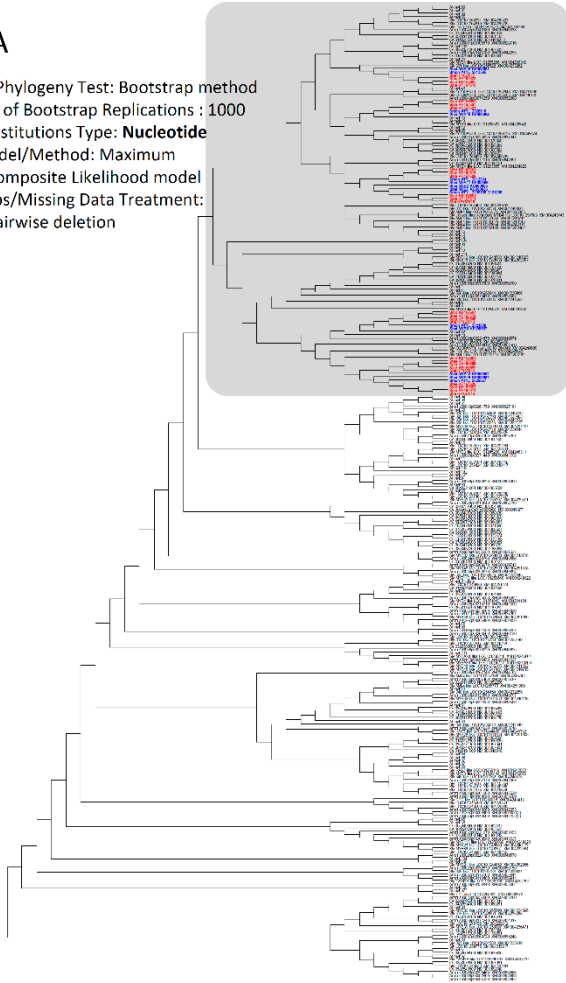


Supplementary Information

A

NJ Phylogeny Test: Bootstrap method
No. of Bootstrap Replications : 1000
Substitutions Type: **Nucleotide**
Model/Method: Maximum
Composite Likelihood model
Gaps/Missing Data Treatment:
Pairwise deletion



B

NJ Phylogeny Test: Bootstrap method
No. of Bootstrap Replications : 1000
Substitutions Type: **Amino acid**
Model/Method: Jones-Taylor-
Thornton (JTT) model
Gaps/Missing Data Treatment:
Pairwise deletion

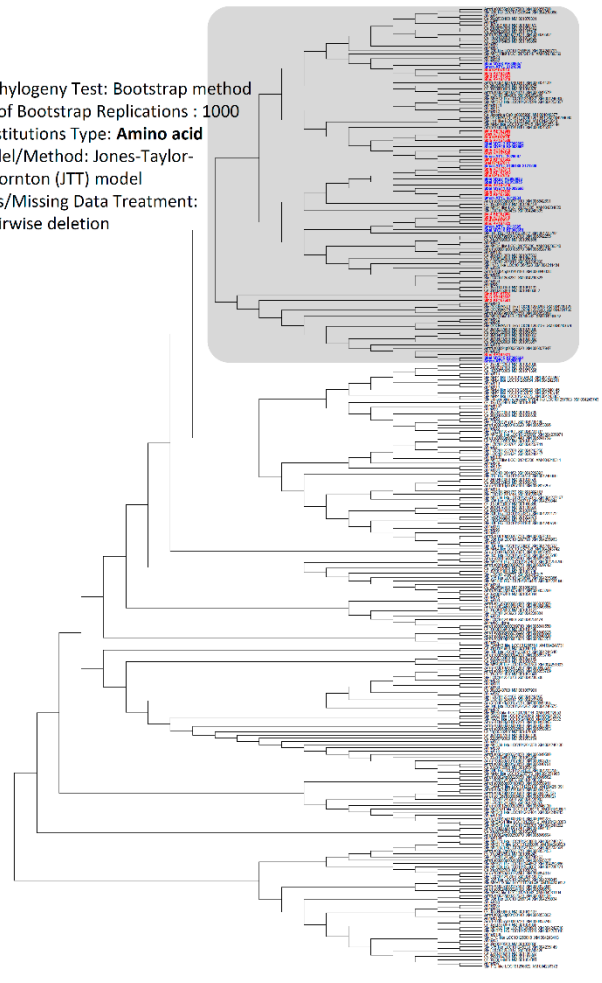


Figure S1. Cont.

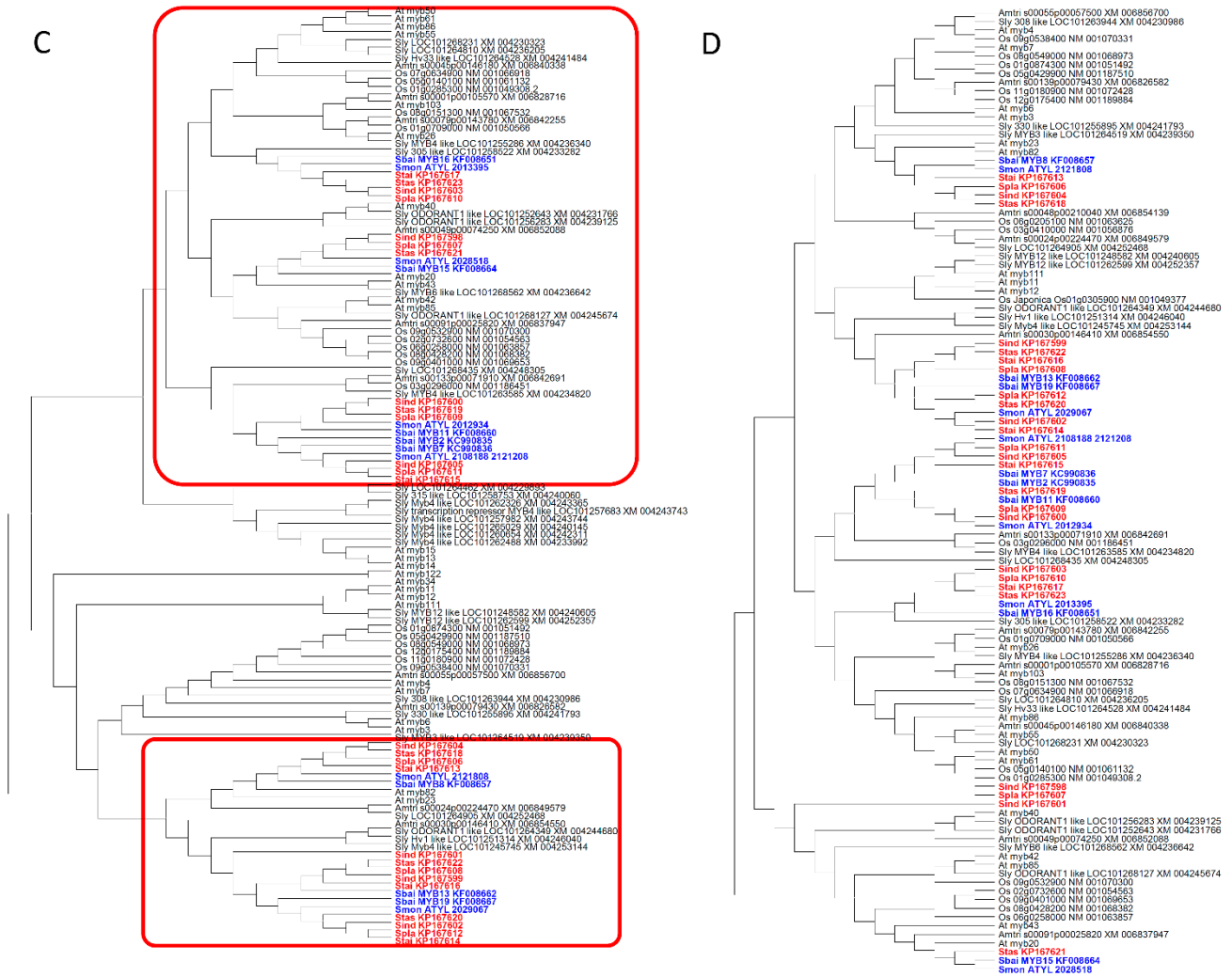


Figure S1. Cladogram of the *Scutellaria* R2R3-MYBs extracted from the inflorescence-bud transcriptomic library of *Scutellaria* and the published R2R3-MYBs from GenBank. (A) Neighbor-joining tree reconstructed based on the nucleotide sequences; (B) Neighbor-joining tree reconstructed based on the amino acid sequences. The gray boxes are the clade including the *Scutellaria* R2R3-MYBs and the corresponding sequences of *Amborella trichopoda* (Amtri), *Arabidopsis thaliana* (At), *Oryza sativa* (Os), and *Solanum lycopersicum* (Sly); (C) and (D) are enlargements of the gray areas of Figure S1A,B, respectively. Based on Figure S1C, the *Scutellaria* R2R3-MYBs and the sequences that were grouped with *Scutellaria* R2R3-MYBs were selected (the red frames) and realigned to re-perform the NJ analysis shown in Figure 1. Sbai: *Scutellaria baicalensis*; Sind: *Scutellaria indica*; Smon: *Scutellaria montana*; Spla: *Scutellaria playfairii*; Stai: *Scutellaria taiwanensis*; Stas: *Scutellaria tashiroi*.

	5	15	25	35	45	55	65	75	85	95	105	
Amtri_s00001p00105570_XM_006828716	MGHHS	CCNQ	KV	RGL	WSP	EDEK	LRYI	T	SN	YG	-CW	SE
Amtri_s00024p00224470_XM_006849579	MGRPC	CCSK	GL	NKG	AW	ED	IL	TD	YK	TH	EG	-KW
Amtri_s00030p00146410_XM_006854550	MGRAP	CCSKV	GL	LR	GP	IA	ED	TL	LK	RY	IE	TH
Amtri_s00045p00146180_XM_006840338	MGRHS	CCYKQ	KL	RK	GL	WSP	EDEK	LRIH	T	SN	YG	-CW
Amtri_s00079p00143780_XM_006842255	MGHHS	CCNQK	KV	RR	GL	WSP	EDEK	LNIH	A	Y	GH	-CW
Amtri_s00133p00071910_XM_006842691	MGRP	PCDDK	SV	RK	GP	WT	ED	AK	IL	AY	GH	-WT
Os_0s01g0285300_NM_001049308.2	MGHHS	CCNQK	KV	RR	GL	WSP	EDEK	LKVI	S	TH	GH	-CW
Os_0s01g0790900_NM_001050566	MGRP	PCCEKV	GL	KK	GP	WT	AD	ED	KL	IN	FI	TH
Os_0s02g0732600_NM_001054563	MGRP	PCCKA	NV	KK	GP	WT	AE	ED	AK	L	AY	TH
Os_0s03g0296000_NM_001186451	MGRHS	CCYKQ	KL	RK	GL	WSP	EDEK	LNIH	T	SN	YG	-CW
Os_0s05g0140100_NM_001061132	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	V	AF	TH
Os_0s07g0634900_NM_001066918	MAK	RAGG	ARK	LR	RL	GL	WSP	EDEK	LNIH	A	Y	GH
Os_0s08g0151300_NM_001067532	MGHHS	CCNQK	KV	RR	GL	WSP	EDEK	LRYI	T	SN	YG	-CW
Os_0s08g0428200_NM_001068382	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
Os_0s09g0401000_NM_001069653	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	V	AF	TH
Os_0s09g0532900_NM_001070300	MGRP	PCCKL	GV	KK	GP	WT	AE	ED	KL	M	SF	TH
At_myb103	MGHHS	CCNQK	KV	RR	GL	WSP	EDEK	LRYI	T	SN	YG	-CW
At_myb20	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
At_myb23	MRM	TRD	KG	E	Y	KK	L	W	T	V	Y	TH
At_myb26	MGHHS	CCNQK	KV	RR	GL	WSP	EDEK	LRYI	T	SN	YG	-CW
At_myb36	MGRAP	CCCKA	NV	KK	GP	WSP	E	D	K	L	I	RF
At_myb42	MGRP	PCCKL	GV	KK	GP	WT	AE	ED	KL	I	RF	TH
At_myb43	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
At_myb5	KK	T	P	CC	T	GM	K	R	GP	W	T	V
At_myb50	MGRHS	CCYKQ	KL	RK	GL	WSP	EDEK	LNIY	T	SN	YG	-CW
At_myb55	MGRHS	CCYKQ	KL	RK	GL	WSP	EDEK	LRYI	T	SN	YG	-CW
At_myb61	MGRHS	CCYKQ	KL	RK	GL	WSP	EDEK	LNIY	T	SN	YG	-CW
At_myb68	MGRAP	CCCKA	NV	KK	GP	WSP	E	D	K	L	I	RF
At_myb82	ME	K	R	E	E	G	K	S	Y	V	R	K
At_myb85	MGRP	PCCKL	GV	KK	GP	WT	AE	ED	KL	I	RF	TH
At_myb86	MGRHS	CCYKQ	KL	RK	GL	WSP	EDEK	LNIY	T	SN	YG	-CW
Sly_L0C1011258522_XM_004233282	-----	Q	V	E	R	K	G	P	W	T	A	E
Sly_L0C1011251314_XM_004246040	MGRAP	CCCKE	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Sly_L0C1011264528_XM_004241484	MGRHS	VFVKE	K	R	K	L	GL	WSP	EDEK	LNIY	T	SN
Sly_L0C1011264810_XM_004236205	MGRHS	CCSVKQ	KL	RK	GL	WSP	EDEK	LNIY	T	SN	YG	-CW
Sly_L0C1011264905_XM_004252468	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
Sly_L0C1011268231_XM_004230323	MGRHS	CCYKQ	KL	RK	GL	WSP	EDEK	LKH	I	T	Y	GH
Sly_L0C1011268435_XM_004248305	MGR	I	P	CC	E	D	NV	K	R	GP	W	T
Sly_L0C1011265745_XM_004253144	MGR	T	PC	CS	K	GM	K	GP	W	T	A	E
Sly_L0C1011265286_XM_004236340	MGRHS	CCNQK	KV	RR	GL	WSP	EDEK	LNIY	T	SN	YG	-CW
Sly_L0C1011263585_XM_004234820	MGRP	PCCKA	NV	KK	GP	WST	ED	AK	IL	AY	TH	GH
Sly_L0C1011268562_XM_004236642	MGRP	PCCKN	EV	KK	GP	WSP	E	D	K	L	I	RF
Sly_L0C1011264349_XM_004246480	MGRAP	CCCKE	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Sly_L0C1011268127_XM_004245674	MGRP	PCCKL	GV	KK	GP	WT	AE	ED	KL	I	RF	TH
Simon_2012934	MGRAP	CCCKT	KV	RK	GP	WSP	E	D	N	L	K	NY
Simon_2013395	AN	S	E	N	G	E	N	G	V	E	N	G
Simon_2028518	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
Simon_2029067	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Simon_2108188_2121208	MGRAP	CCCKA	NV	KK	GP	WSP	E	D	A	Q	L	K
Simon_2121808	ME	K	K	G	-----	-	V	K	R	G	S	W
Sbai_MYB11_KF008660	MGRAP	CCCKT	KV	RK	GP	WSP	E	D	N	L	K	NY
Sbai_MYB13_KF008662	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Sbai_MYB15_KF008664	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
Sbai_MYB16_KF008665	AN	S	G	N	G	I	G	E	S	N	G	V
Sbai_MYB19_KF008667	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Sbai_MYB2_KC990835	MGRAP	CCCKG	NV	KK	GP	WSP	E	D	A	K	L	I
Sbai_MYB7_KC990836	MGRAP	CCCKA	NV	KK	GP	WSP	E	D	A	Q	L	K
Sbai_MYB8_KF008667	ME	K	K	G	V	-----	-	V	K	R	G	A
Sind_16195	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
Sind_17590	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Sind_18942	MGRAP	CCCKT	KV	RK	GP	WSP	E	D	N	L	K	NY
Sind_21139	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
Sind_21139_2	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Sind_28842	AN	S	E	N	G	E	N	G	V	E	N	G
Sind_31679	ME	K	R	A	E	-----	-	V	K	R	G	A
Sind_comp14321	MGRAP	CCCKA	NV	KK	GP	WSP	E	D	A	Q	L	K
Spla_10351	ME	N	R	A	E	-----	-	V	K	R	G	A
Spla_13224	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
Spla_2011	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Spla_25860_32398	MGRAP	CCCKT	KV	RK	GP	WSP	E	D	N	L	K	NY
Spla_28842	AN	S	E	N	G	E	N	G	V	E	N	G
Spla_29147	MGRAP	CCCKA	NV	KK	GP	WSP	E	D	A	Q	L	K
Spla_3616	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Stai_12687	ME	K	K	A	E	-----	-	V	K	R	G	A
Stai_17601	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Stai_27949	MGRAP	CCCKA	NV	KK	GP	WSP	E	D	A	Q	L	K
Stai_6424	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Stai_7654	AN	S	E	N	G	E	N	G	V	E	N	G
Stas_10585	ME	K	R	A	E	-----	-	V	K	R	G	A
Stas_1129	MGRAP	CCCKT	KV	RK	GP	WSP	E	D	N	L	K	NY
Stas_12182	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Stas_14132	MGRP	PCCKV	GL	KK	GP	WT	AE	ED	KL	I	RF	TH
Stas_4055	MGRS	PCCKV	GL	RK	GP	WST	ED	LL	T	N	Y	IN
Stas_9284	AN	S	E	N	G	E	N	G	V	E	N	G

Figure S2. Amino acid alignments of the conserved region of the R2R3 Myb genes used for the NJ tree reconstruction in Figure 1. Amtri: *Amborella trichopoda*; At: *Arabidopsis thaliana*; Os: *Oryza sativa*; Sbai: *Scutellaria baicalensis*; Sind: *Scutellaria indica*; Sly: *Solanum lycopersicum*; Smon: *Scutellaria montana*; Spla: *Scutellaria playfairii*; Stai: *Scutellaria taiwanensis*; Stas: *Scutellaria tashiroi*.