A

```
1    ATGGATGTTAAGAAAGGTGGGTCTGTAGTACAAGCACAAGTGAAGTTGCAGAAGCATAAC
     M  D  V  K  K  G  G  S  V  V  Q  A  Q  V  K  L  Q  K  H  N
61   GAAAAGGAGATGGGCATGAGAAAAGGTCCATGGGCGGTTGAGGAGGACACCATTCTGGTC
     E  K  E  M  G  M  R  K  G  P  W  A  V  E  E  D  T  I  L  V
121  AATTACATCGCCACACACGGTGAAGGCCACTGGAATTCCGTGGCACGATGTGCAGGTCTA
     N  Y  I  A  T  H  G  E  G  H  W  N  S  V  A  R  C  A  G  L
181  AGGAGGAGTGGGAAGAGTTGCAGATTAAGGTGGCTAAACTACTTGCGCCCAGACGTGCGG
     R  R  S  G  K  S  C  R  L  R  W  L  N  Y  L  R  P  D  V  R
241  CGTGGAAATATCACACTCCAAGAACAAATATTAATTCTCGACCTTCACTCTCGCTGGGGC
     R  G  N  I  T  L  Q  E  Q  I  L  I  L  D  L  H  S  R  W  G
301  AACAGGTGGTCAAAGATTGCTCAACAGCTGCCAGGAAGGAACAGACAACGAAATAAAGAAC
     N  R  W  S  K  I  A  Q  Q  L  P  G  R  T  D  N  E  I  K  N
361  TATTGGAGGACCAGAGTGATAAAACAAGCGAAGCAGCTAAAGTGCGATGTGAATAGCAAA
     Y  W  R  T  R  V  I  K  Q  A  K  Q  L  K  C  D  V  N  S  K
421  CAGTTCAGAGACACGTTGCGTTACGTTTGGATGCCGCGCTTGCTGGAGCGGCTTCAGCCC
     Q  F  R  D  T  L  R  Y  V  W  M  P  R  L  L  E  R  L  Q  P
481  ACATCACAAGCACTGGAACCCAATCAAAGTGGACTTGTTGTTACACGCTTCATCATCACTG
     T  S  Q  A  L  E  P  N  Q  S  G  L  V  L  H  A  S  S  S  L
541  CTTCCTTCGAATTCCGACCATAGTATTGAAAGGGGGTCGGATCTGTGGCCAGGTTTCAAT
     L  P  S  N  S  D  H  S  I  E  R  G  S  D  L  W  P  G  F  N
601  AACCAAATGTTGTTGGAACAGGGGAGTGGCGGTGACTTGTTGGAAAGTTTGTGGGATGAC
     N  Q  M  L  L  E  Q  G  S  G  G  D  L  L  E  S  L  W  D  D
661  GACAATATGTGCTTTTTGCAACAGCTTTCTTATGACCTCCAAATGAAATAA
     D  N  M  C  F  L  Q  Q  L  S  Y  D  L  Q  M  K  *
```
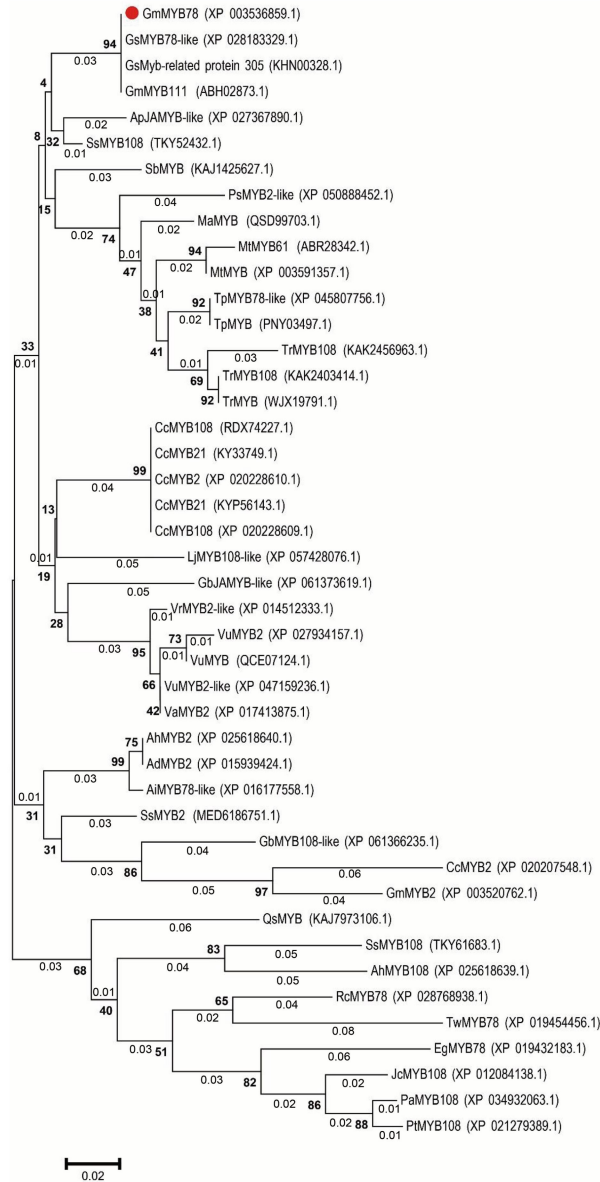
— MYB domain

— MYB domain

B

GmMYB78 (XP 003536859.1)
GsMYB78-like (XP 028183329.1)
GsMyb-related protein 305 (KHN00328.1)
GmMYB111 (ABH02873.1)
ApJAMYB-like (XP 027367890.1)
SsMYB108 (TKY52432.1)
SbMYB (KAJ1425627.1)
PsMYB2-like (XP 050888452.1)
MaMYB (QSD99703.1)
MtMYB61 (ABR28342.1)
MtMYB (XP 003591357.1)
TpMYB78-like (XP 045807756.1)
TpMYB (PNY03497.1)
TrMYB108 (KAK2456963.1)
TrMYB108 (KAK2403414.1)
TrMYB (WJX19791.1)
CcMYB108 (RDX74227.1)
CcMYB21 (KY33749.1)
CcMYB2 (XP 020228610.1)
CcMYB21 (KYP56143.1)
CcMYB108 (XP 020228609.1)
LjMYB108-like (XP 057428076.1)
GbJAMYB-like (XP 061373619.1)
VrMYB2-like (XP 014512333.1)
VuMYB2 (XP 027934157.1)
VuMYB (QCE07124.1)
VuMYB2-like (XP 047159236.1)
VaMYB2 (XP 017413875.1)
AhMYB2 (XP 025618640.1)
AdMYB2 (XP 015939424.1)
AiMYB78-like (XP 016177558.1)
SsMYB2 (MED6186751.1)
GbMYB108-like (XP 061366235.1)
CcMYB2 (XP 020207548.1)
GmMYB2 (XP 003520762.1)
QsMYB (KAJ7973106.1)
SsMYB108 (TKY61683.1)
AhMYB108 (XP 025618639.1)
RcMYB78 (XP 028768938.1)
TwMYB78 (XP 019454456.1)
EgMYB78 (XP 019432183.1)
JcMYB108 (XP 012084138.1)
PaMYB108 (XP 034932063.1)
PtMYB108 (XP 021279389.1)

0.02

Figure S1. Nucleotide and amino acid sequences analyses of GmMYB78. (**A**)The MYB domain is shown in shadow. (**B**) Phylogenetic analysis of GmMYB78 with homologous genes.
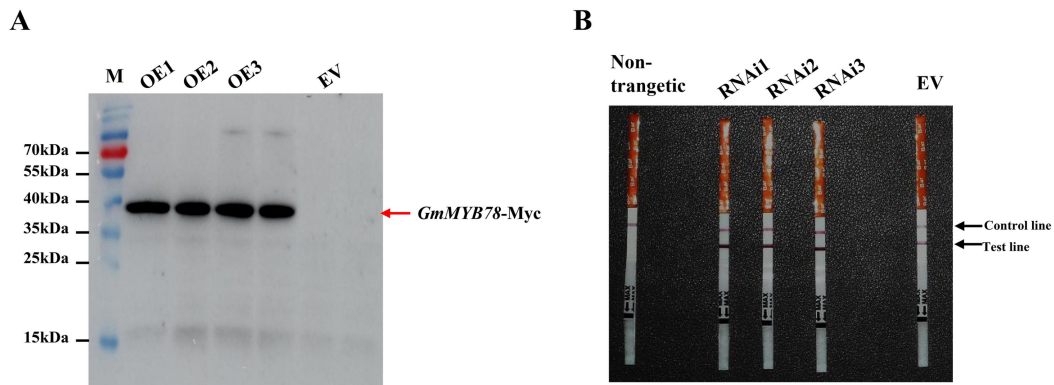
Figure S2. Analyses of *GmMYB78* transgenic soybean hairy roots. (**A**) Immunoblots showing the expression of the GmMYB78-Myc fusion protein in *GmMYB78*-OE transgenic hairy roots and the EV controls. The total protein extracts were analyzed using a 12% SDS-PAGE, and the immunoblot was probed with anti-Myc antibody. (**B**) The *GmMYB78*-RNAi and EV transgenic hairy roots were tested using QuickStix Kit for LibertyLink (bar) strips.
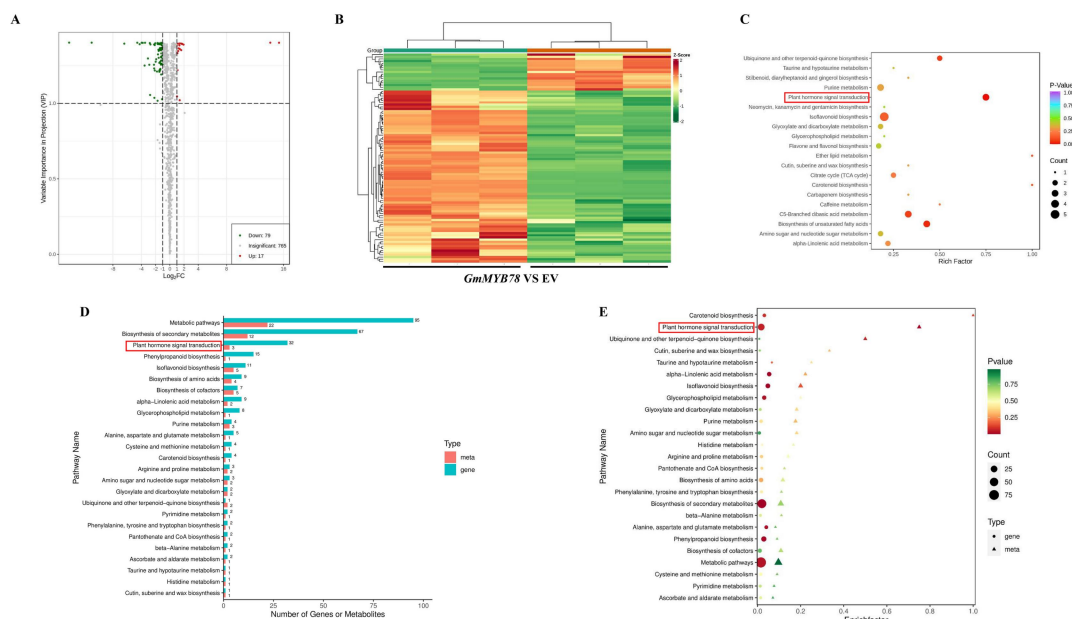

Figure S3. Metabolome analysis of metabolites in *GmMYB78* overexpression transgenic soybean hairy roots. (**A**) Volcano plots of significantly differentially metabolites in *GmMYB78*-OE vs. EV transgenic soybean hairy roots. (**B**) Heat map of significantly differentially metabolites between the EV and *GmMYB78*-OE transgenic soybean hairy roots. Using VIP ≥ 1 and fold-change set at ≥2 as the screening criteria, a total of 96 differentially metabolites were identified. (**C**) Metabolite annotation of KEGG classification of differentially metabolites. The differentially expressed metabolites are mainly involved in multiple biological processes, including plant hormone signal transduction. (**D**) KEGG classification bar chart of combined analysis of transcriptome and metabolome. The bar chart shows the number of differential metabolites and differential genes enriched to a certain pathway. (**E**) KEGG classification bubble chart of combined analysis of transcriptome and metabolome. Showing the top 25 pathways for P-value.
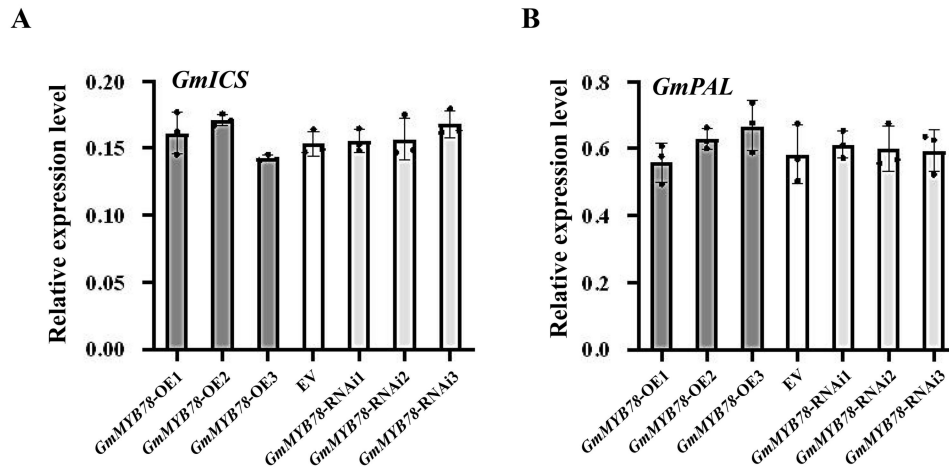
Figure S4. Investigation of the relationship between *GmMYB78* and the salicylic acid (SA) pathway in soybean. (**A,B**) Relative transcript level of *GmICS* (A) or *GmPAL* (B) in *GmMYB78*-OE, *GmMYB78*-RNAi and EV transgenic soybean hairy roots. The level of the control sample (EV lines) was set to unity. The reference gene *GmEF1β* and *GmTUB4* were used as reference genes. The experiment was performed on three biological replicates, each with three technical replicates, and was statistically analyzed using Student's *t*-test (*$P < 0.05$, **$P < 0.01$). Bars indicate the standard error of the mean.
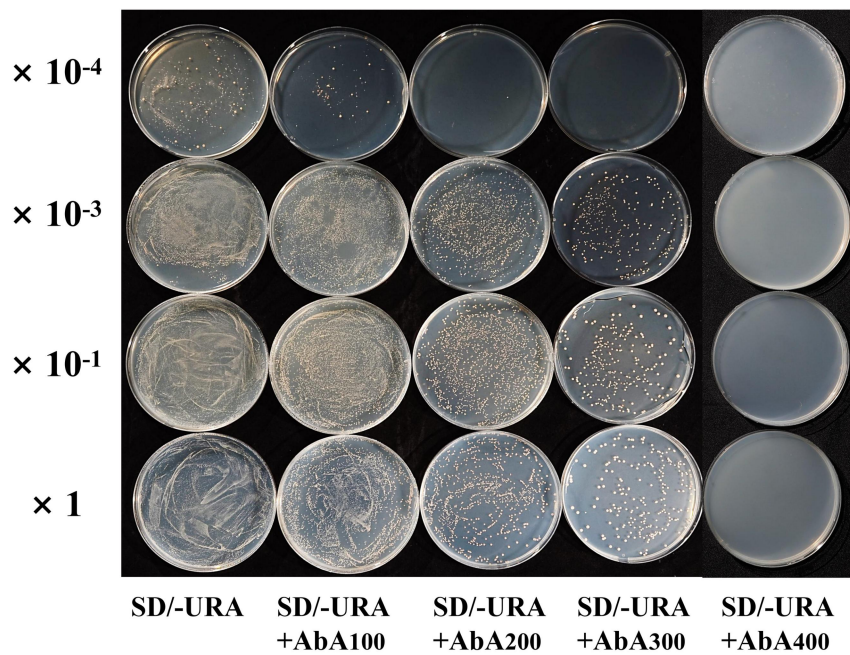

Figure S5. Screening of AbA concentration of Y1H [pBait-AbAi-*pGmbZIP25*] strain. ×1 represents the OD$_{600}$ of the bait strain Y1H[pBait-AbAi-*pGmMYB78*] as 0.02; 10^{-1}, 10^{-2} or 10^{-3} represent dilutions of 10

times, 100 times, 1000 times, respectively. They were spread on solid SD/-Ura (SD medium without Ura) medium with different concentrations of AbA (100 ng/mL, 200 ng/mL, 300 ng/mL, 400 ng/mL) and incubated inverted on plates for 5 days.
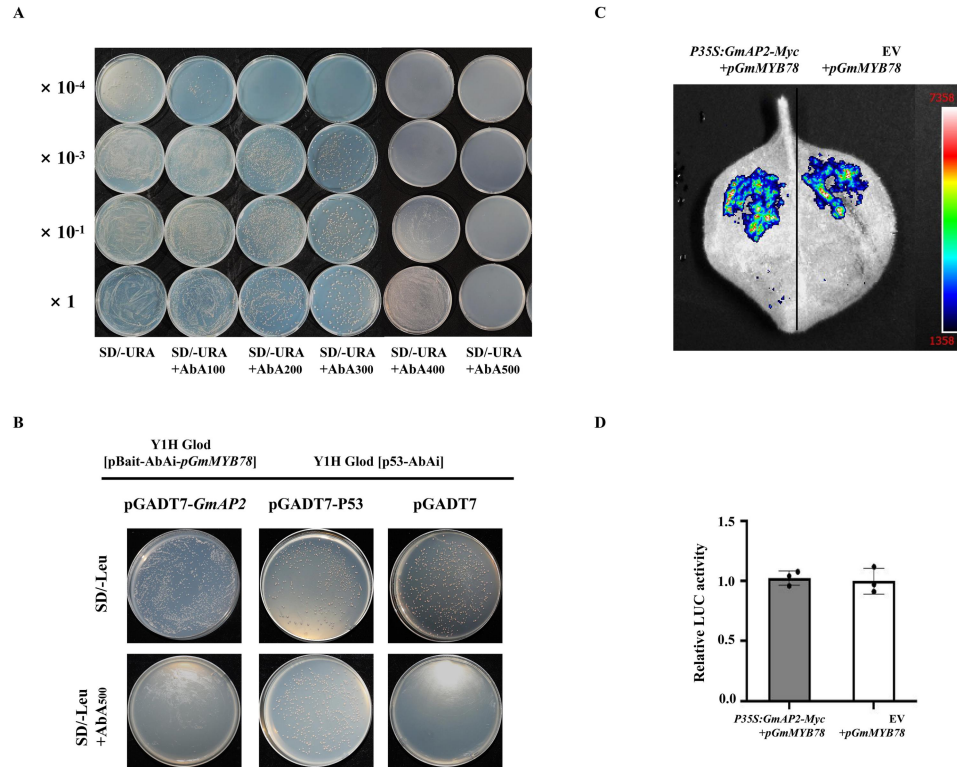

Figure S6. GmAP2 cannot directly regulate *GmMYB78*. (**A**) Screening of AbA concentration of Y1H [pBait-AbAi-*pGmMYB78*] strain. ×1 represents the OD_{600} of the bait strain Y1H[pBait-AbAi-*pGmMYB78*] as 0.02; 10^{-1}, 10^{-2} or 10^{-3} represent dilutions of 10 times, 100 times, 1000 times, respectively. They were spread on solid SD/-Ura (SD medium without Ura) medium with different concentrations of AbA (100 ng/mL, 200 ng/mL, 300 ng/mL, 400 ng/mL, 500 ng/mL) and incubated inverted on plates for 5 days. (**B**) Verification of GmAP2 regulate *GmMYB78* in yeast. (**C**) Dual-luciferase assay in *N. benthamiana* leaves showing that GmAP2 promotes the expression of *GmMYB78*. Representative photographs are shown. (**D**) Detection of LUC/Rluc activity to verify whether GmAP2 promotes the expression of *GmMYB78*. The combination of the reporter construct (*pGmMYB78*: LUC) and EV was used as a control. The experiment was performed on three biological replicates, each with three technical replicates, and the results were statistically analyzed using Student's *t*-test (*$P < 0.05$, **$P < 0.01$). Bars indicate the standard deviation of the mean (n=3).

Table S1. List of upstream protein of GmMYB78.

Library screening results by yeast one-hybrid

GenBank ID	Gene	Number of clones
XM-003519791.5	Zine-finger homeodomain protein	17
XM-014770182.3	Peroxisomal biogenesis factor 11family protein	1
XM-003522700.5	AT-hook motif nuclear-localized protein 17	2
XM-003543501.5	Transcription factor GmbHLH122	1

Table S2. List of primers used in this study.

Primers for quantitative real-time PCR

Primer name	Primers (5′-3′)
GmMYB78-RT-qPCR	F: GGAGGAGTGGGAAGAGTTGC
	R: GTGATATTTCCACGCCGCAC
GmTUB4- RT-qPCR	F: GGCGTCCACATTCATTGGA
	R: CCGGTGTACCAATGCAAGAA
GmEF1β- RT-qPCR	F: CCACTGCTGAAGAAGATGATGATG
	R: AAGGACAGAAGACTTGCCACTC
TEF1- RT-qPCR	F: TGATCGTGCTGAACCACCC
	R: CGAGCGACGGTCCATCTT
PSPEL1- RT- qPCR	F: CCGCGT ACG TGGC TT TGGTGAG
	R: ATCT TGGCGAC TGAGGC TGC T TAC
GmPR1- RT-qPCR	F: GCCCTTATCGGGGTGTTTCT
	R: TCCTTTCTTCCTGTGTGTATTGC
GmPR2- RT-qPCR	F: CTTCCGTATGTGTTGCTTCGT
	R: TTTCCTGTATTGTGCCGCCTT
GmPR3- RT-qPCR	F: TTCTCCTTCTCCTCTTCTCCTT
	R: TGCCTTTGCCTACTGTATGT
GmPR10- RT-qPCR	F: TTCGCCGTGTGTGCTTATTGT
	R: TATCCTCCTCTTCCTCTATTGTC
GmAOS1- RT-qPCR	F: AAGTGCATCACTGACGAGGG
	R: TGTGGGAACCCACCTTCAAC
GmJAZ1- RT-qPCR	F: AGTGGTGTGGCTCAGGAATG
	R: GAGGGGCAAGGGCTATGTA
GmJAZ2- RT-qPCR	F: GCAGAAGAACAACAACGACGG
	R: CGAGGGGGTTTGACGTAGAG
GmbHLH122- RT-qPCR	F: TGCCACTCACCCAAGAA
	R: CCACAGCCAAGTCCAAC

Primers for gene cloning

Primer name	Primers (5′-3′)
GmMYB78	F: ATGGATGTTAAGAAAGGTGGGT
	R: TTATTTCATTTGGAGGTCATAAGA
GmbHLH122	F: ATGGAGTCAGATCTCGAGCAAC
	R: TTATTGTTGCTTCTCGTGTGA

Primers for constructs in transformation of soybean hairy roots

Primer name	Primers (5′-3′)
GmMYB78-Myc	F:CGGGGGACTCTTGACCATGGAAATGGATGTTAAGAAAGGTGGGT

	R: CCCACGTGGCTAGC AGATCT TTTCATTTGGAGGTCATAA
GmMYB78 RNAi1	F:TTTCATTTGGAGAGGACACGCTCGAGATGGATGTTAAGAAAGGTGGGTC
	R:AATCATCGATTGGGCGCGCCCCATGGTTCGCTTGTTTTATCACTCTGGT
GmMYB78 RNAi2	F:TCCCGGGTCTTAATTAACTCTCTAGAATGGATGTTAAGAAAGGTGGGTC
	R:GTCAATTTGCAGGTATTTGGATCCTTCGCTTGTTTTATCACTCTGGT

Primers for constructs in yeast one-hybrid assays

Primer name	Primers (5′-3′)
pGmMYB78	F: TTAACATATTAGATTGTGTAAAA
	R: ACTTATTGGGGGCATGTGGAGTT
pGmMYB78-pBait-AbAi	F: AAAAGCTTGAATTCGAGCTCTTAACATATTAGATTGTGT
	R: TACAGAGCACATGCCTCGAGACTTATTGGGGGCATGTG
GmbHLH122-AD	F: CCATGGAGGCCAGTGAATTCATGGAGTCAGATCTCG
	R: AGCTCGAGCTCGATGGATCCTTATTGTTGCTTCTCGT
pGmbZIP25	F: CTATTTGATCCATACACATA
	R: AGAAGGAGGAAGGAAAGAG
pGmbZIP25-pBait-AbAi	F: AAAAGCTTGAATTCGAGCTC CTGTTGTCTATCAACGGATAAA
	R:TACAGAGCACATGCCTCGAGAAGCTGATTCTCTGTTGCTGTTG

Primers for subcellular localization

Primer name	Primers (5′-3′)
GmMYB78-1302	F: CGGGGGACTCTTGACCATGGATGTTAAGAAAGGTGGGT
	R: AGATCTGACTAGT TTATTTCATTTGGAGGTCATAAGA

Primers for transient expression assay

Primer name	Primers (5′-3′)
GmMYB78-BD	F: TGGCCATGGAGGCCGAATTCATGGATGTTAAGAAAGGTGG
	R: CGCTGCAGGTCGACGGATCCTTATTTCATTTGGAGGTCATAA

Primers for transient transcription dual-luciferase assay

Primer name	Primers (5′-3′)
pGmMYB78-LUC	F: GTCGACGGTATCGATAAGCTTTTAACATATTAGATTGTGT
	R: CGCTCTAGAACTAGTGGATCCACTTATTGGGGGCATGTG
GmbHLH122-3301	F: CGGGGGACTCTTGA CCATGGAAATGGAGTCAGATCTCG
pGmbZIP25-LUC	R: CCCACGTGGCTAGC AGATCTTTGTTGCTTCTCGT
	F: GTCGACGGTATCGATAAGCTTCTATTTGATCCATACACATA
	R: CGCTCTAGAACTAGTGGATCCAGAAGGAGGAAGGAAAGAG

Primers for transcriptome sequencing analysis

Primer name	Primers (5′-3′)
GmPRR13 -qPCR	F: AAGTGCATCACTGACGAGGG
	R: TGTGGGAACCCACCTTCAAC

GmbZIP25-qPCR	F: AGTGGTGTGGCTCAGGAATG
	R: GAGGGGCAAGGGCTATGTA
GmMYB177-qPCR	F: GATGCTTCCCTTTTGGCTCCTT
	R: AACAACGGAGCCGTGCAAA
GmPR-4A -qPCR	F: ACCCTGAGCAGATTGGTTGG
	R: ATCCTCACCGTAGCCTCAGT
GmLOX9-qPCR	F: GTCCATCAGTATTGGTGTAG
	R: GTCCATCAGTCTTGGTGGCA
GmCHS13-qPCR	F: GAAGCTGGGAAAACAAGCCG
	R: GAGAAGCTTCGTGAGCTGGT
GmBVT -qPCR	F: TGGCCCCGGAACCATTAAA
	R: CTCCACTGTGTCTGGCAACC
GmARM -qPCR	F: ACTAGCACTGGTGGTGTTGG
	R: TGCCGTCATCCAAGTTTCCA