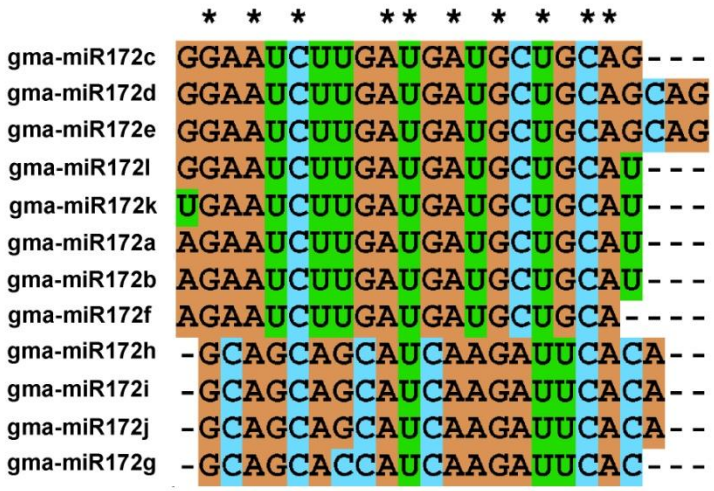
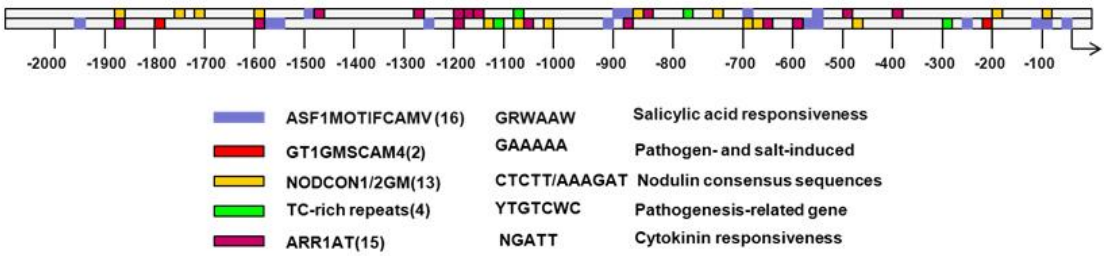
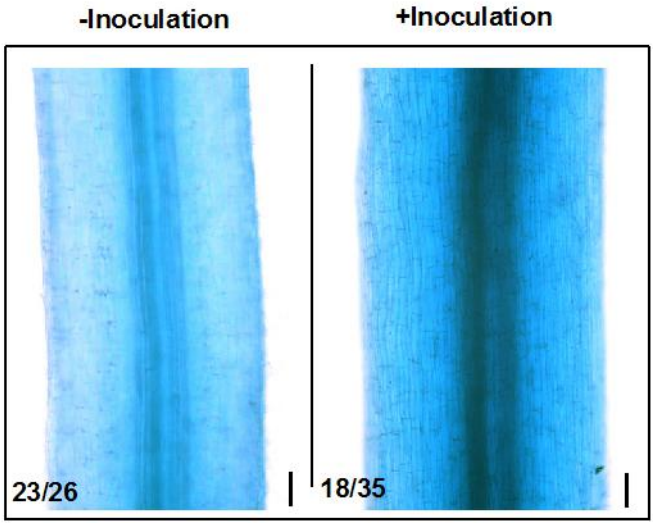


Supplemental Figure 1. Analysis of mature miRNA sequences of the miR172 family members.

Mature sequences of miR172 family members were aligned using software MEGA5. The stars indicate nucleotides conserved in all of the miR172 family members.

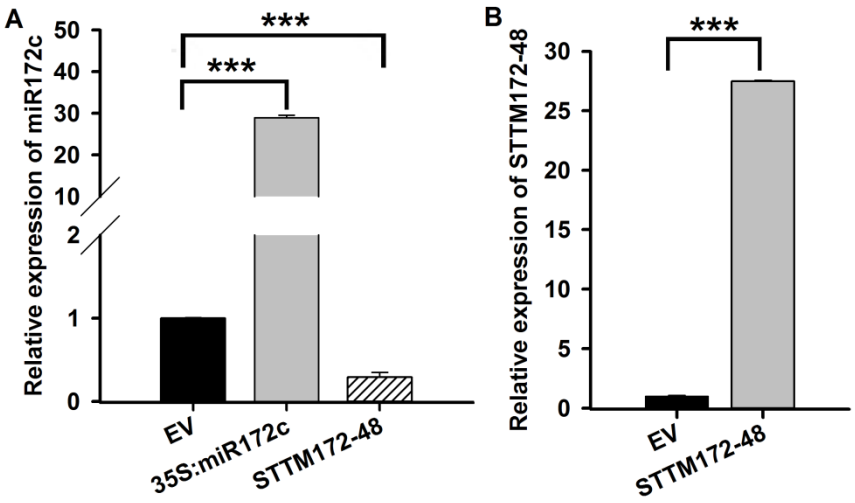


Supplemental Figure 2. Analysis of the miR172c promoter.
The promoter sequences (2000 bp) upstream of pre-miR172c (http://www.phytozome.net/) was chosen for *cis*-element analysis using software at http://www.dna.affrc.go.jp/ PLACE/.

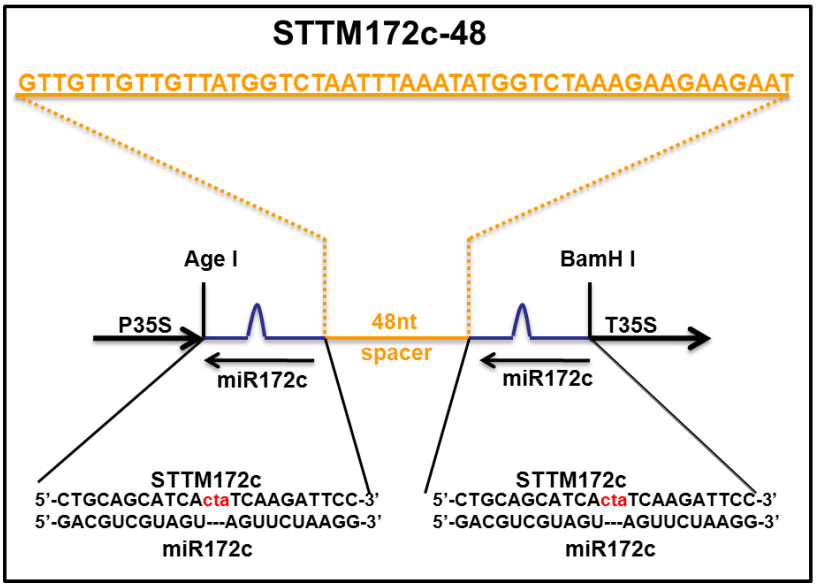


Supplemental Figure 3. Histochemical analysis of miR172c.

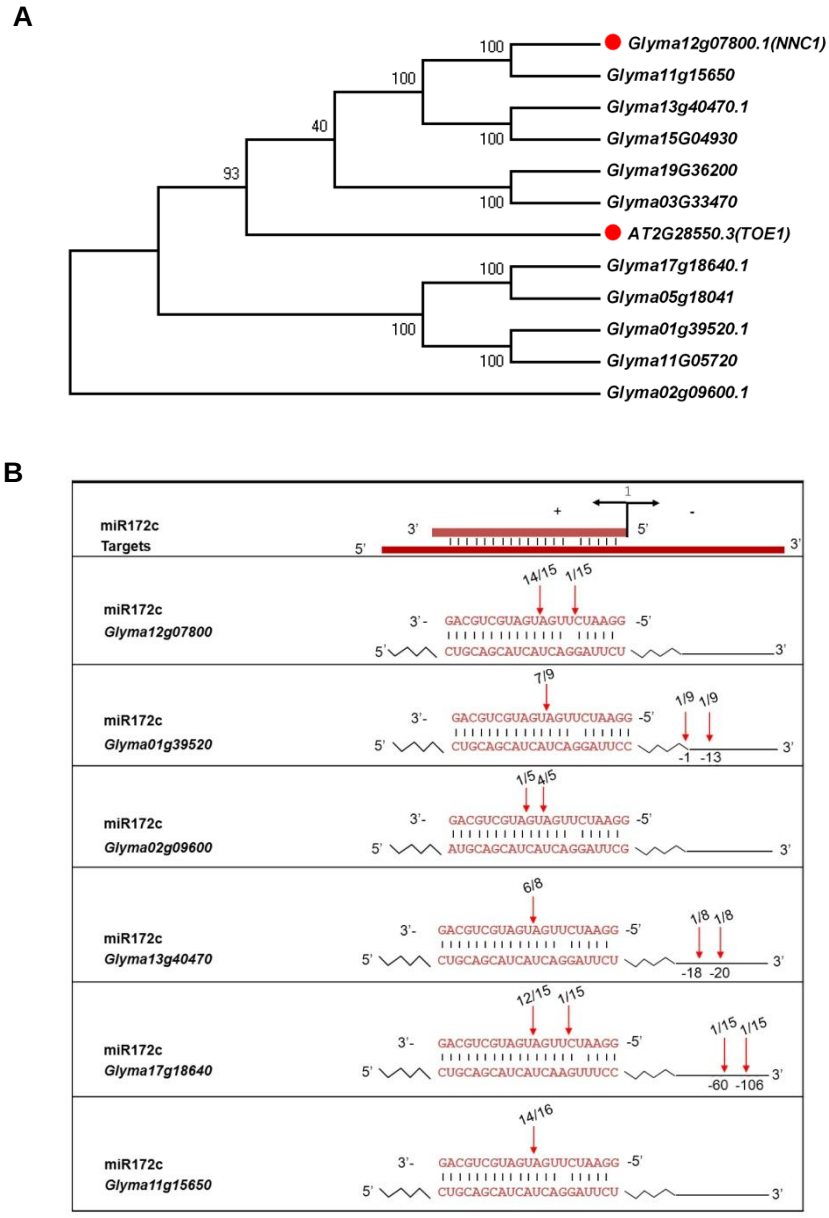
The construct harboring *promiR172c:GUS* was transformed using the hairy root transformation system. The chimeric transgenic plants were then inoculated with *B. japonicum*. The positive transgenic roots at 10 days after inoculation were analyzed using GUS assay. The number of independent transgenic roots showing the representative staining pattern out of the number of roots examined is indicated in each panel. Bar=100 μm. –Inoculation: Without inoculation; + Inoculation: inoculated with rhizobia.



Supplemental Figure 4. qRT-PCR analysis of the miR172c transcript level.
(A) qRT-PCR analysis of the miR172c transcript level in transgenic roots expressing the empty vector (EV), *35S:miR172c*, and STTM172-48. The expression levels were normalized to that of soybean miR1520d. **(B)** qRT-PCR analysis of the STTM172-48 level in the transgenic hairy roots. The expression levels were normalized against the geometric mean of the soybean gene *ELF1b*. Expression levels shown are means ± standard errors from three replicates. Student's *t*-test was performed; statistically significant results are marked with '***' (P<0.001).

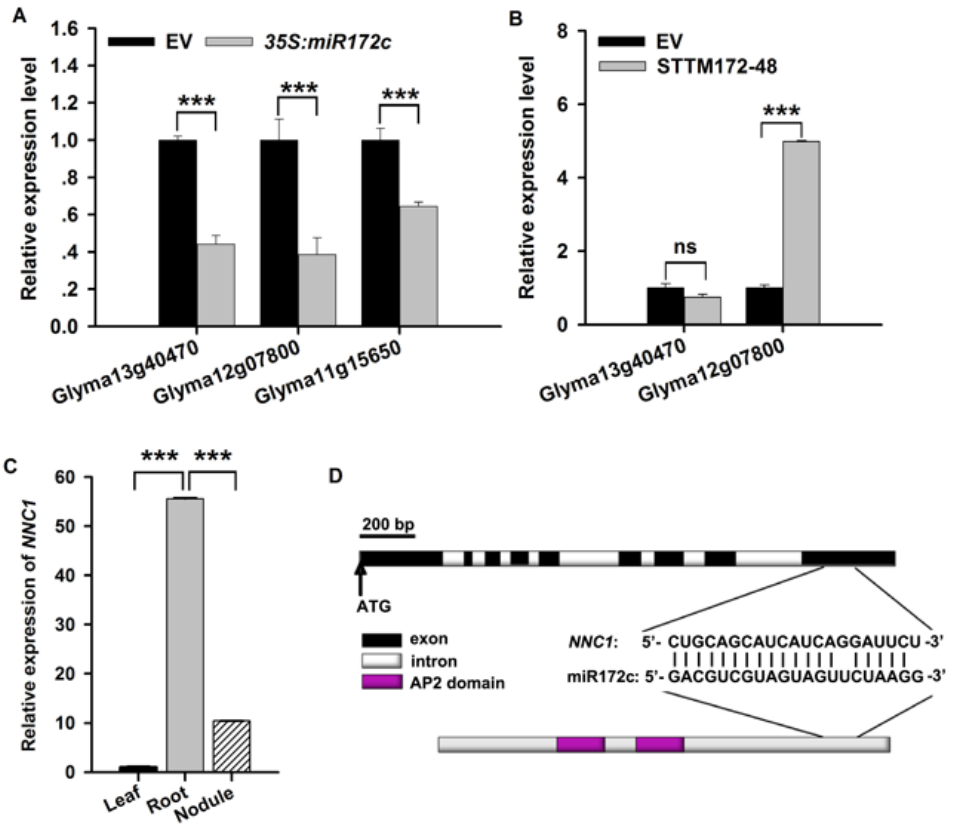


Supplemental Figure 5. Diagram of STTM172c-48 structure showing the design strategy. The fragment and sequence in range indicate the spacer region and the spacer sequence (48 bp). The letters in red indicate the bulge sequences in the miRNA binding sites.



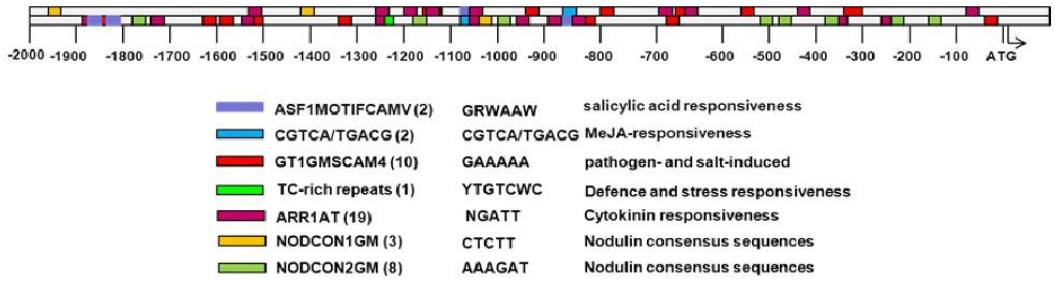
Supplemental Figure 6 Prediction and experimental validation of the target genes of miR172c.

(A) Phylogenetic analysis of 11 predicted target genes of miR172c in soybean. The phylogenetic tree was constructed using MEGA5 phylogenetic analysis software. **(B)** Experimental validation of miR172c target genes and the cleavage sites using 5'RACE. Vertical arrows indicate the 5' termini of the miRNA-guided cleavage products, as identified by 5'-RACE, with the frequency of clones shown.



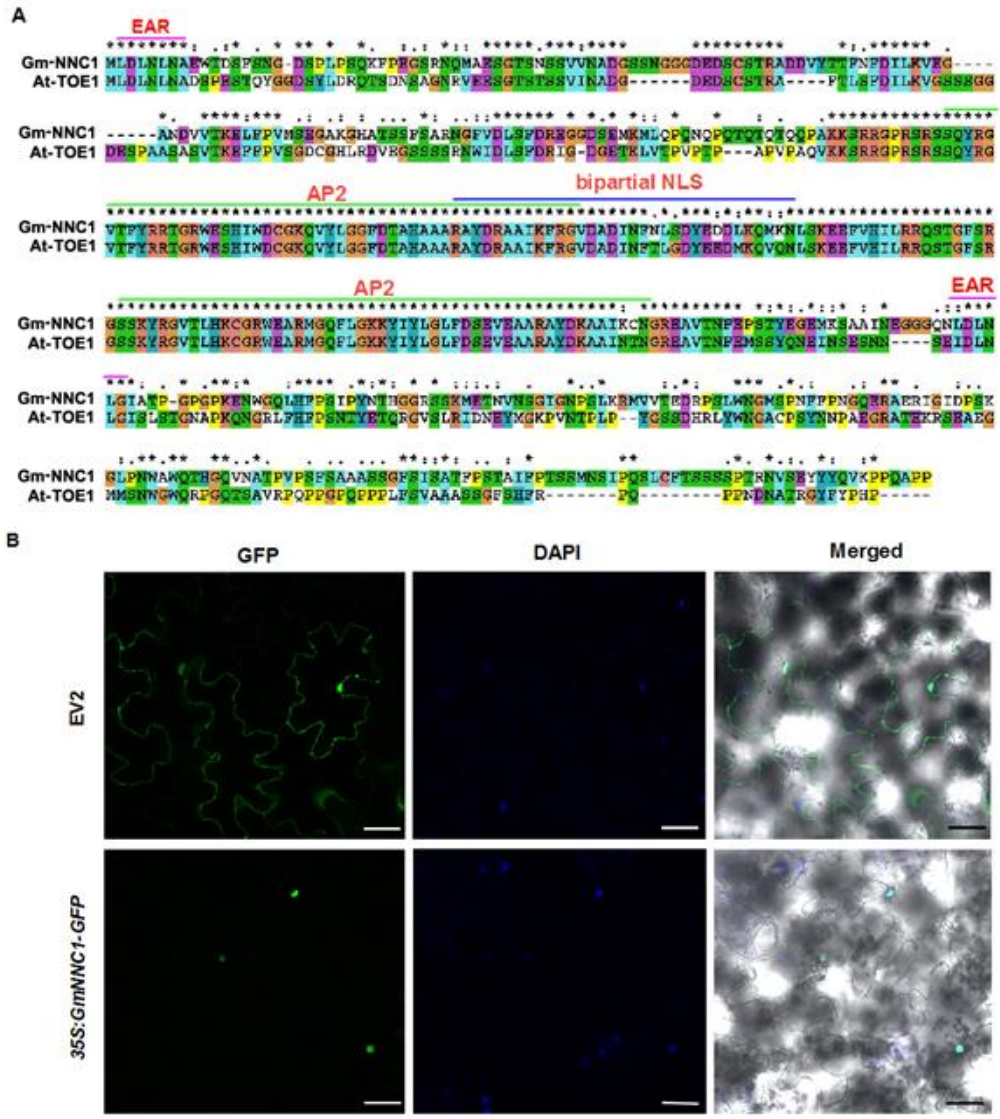
Supplemental Figure 7. Gene expression analysis of target genes and structures of *NNC1.*

(A) qRT-PCR analysis of the transcript levels of *glyma13g40470, glyma12g07800* and *glyma11g15650* in the transgenic roots transformed with the empty vector and *35S:miR172c* at 28 days after inoculation. **(B)** qRT-PCR analysis of the transcript levels of *glyma13g40470* and *glyma12g07800* in the transgenic roots transformed with the empty vector and *35S:STTM-miR172-48* (STTM172-48) at 6 days after inoculation. **(C)** qRT-PCR analysis of the transcript levels of *NNC1* in the leaf, root and nodule at 28 days after inoculation. The expression levels were normalized against the geometric mean of the soybean reference gene *GmELF1b*. Expression levels shown are means ± standard errors from three replicates. The Student's *t*-test was performed and the significantly different treatments were marked with '***' (P<0.001) and '**' (P<0.01). ns: not significant (P>0.05). **(D)** *NNC1* gene contains 9 exons and 8 introns with the miR172c cleavage site in 9th exon.



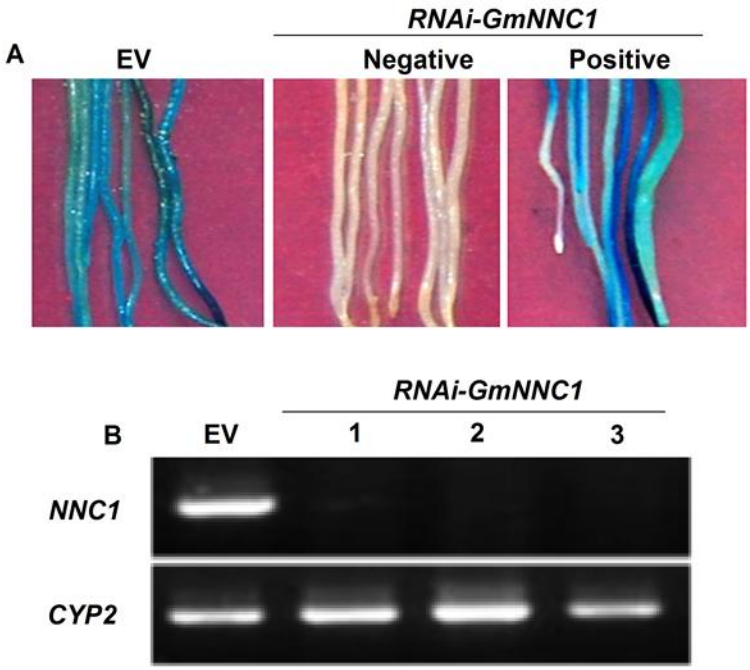
Supplemental Figure 8. Promoter analysis of *NNC1.*

The 2000 bp sequences upstream of start codon 'ATG' of the *NNC1* were chosen as promoter sequences (http://www.phytozome.net/) for analyzing *cis*-elements using software at http://www.dna.affrc.go.jp/PLACE/.

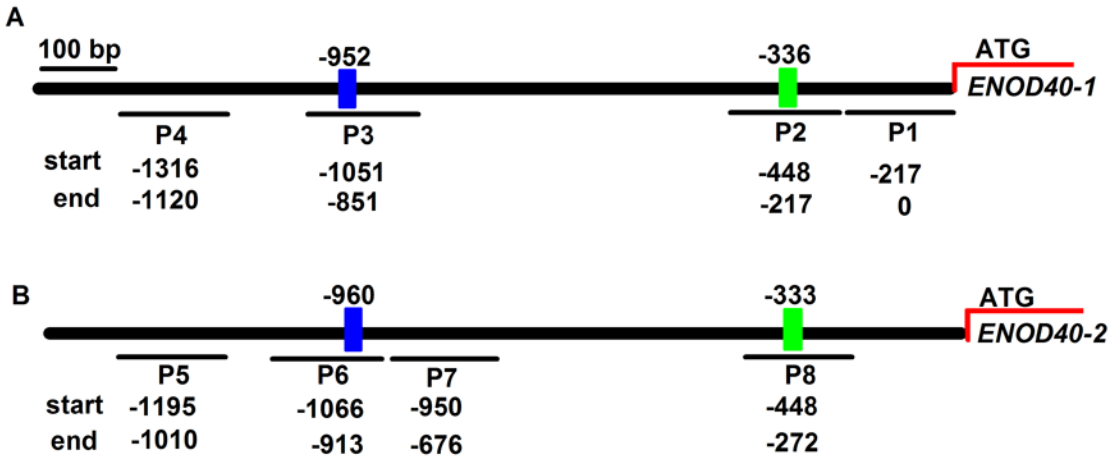


Supplemental Figure 9. NNC1 protein analysis.

(A) Soybean *NNC1* encodes a putative AP2 family protein with high similarity to *Arabidopsis* TOE1. The amino acid sequence of Gm-NNC1 and At-TOE1 were aligned using MEGA5 software. The green line indicates the AP2 domain. The purple line indicates the EAR motif. The blue line indicates the bipartial NLS motif, which is analyzed in the http://rostlab.org/owiki/index.php/predict. **(B)** NNC1 is localized in nucleus. The *35S:GmNNC1* constructs were transformed in the tobacco leaves and the expression of NNC1 was imaged under confocal microscopy. DAPI, 4',6-diamidino-2- phenylindole. Bar=40 μm.



Supplemental Figure 10. Identification of *NNC1* RNAi knockdown roots.
(A) Identification of transgenic hairy roots harboring *RNAi-GmNNC1* and the empty vector by GUS assay. One-week-old hairy roots about 1-1.5 cm in length grown in root induction medium were collected and stained in GUS staining solution. The GUS-positive lines were used for gene expression analysis, rhizobium inoculation and nodulation evaluation. **(B)** RT-PCR analysis of *NNC1* expression in the GUS-positive roots transformed with *RNAi-GmNNC1*construct and the empty vector. After nodule numbers were counted, segments (2 cm in length) of *RNAi-GmNNC1* transgenic roots were used for expression analysis of *NNC1*. The content of *NNC1* mRNA was normalized with respect to *CYP2* mRNA.



Supplemental Figure 11. Analysis of *ENOD40* promoters.
The promoters of *ENOD40-1* (1718 bp) **(A)** and *ENOD40-2* (1496 bp) **(B)** were analyzed. Both CCTCGT and TTAAGGTT (showing in blue and green respectively) binding sites can be observed in promoters of *ENOD40-1* and *ENOD40-2*. P1-P8 indicates eight DNA fragments in these two promoters.

Supplemental Table 1. The number of miR172 family members in plants.

Classes	Families	Species	Members
Dicotyledon	Brassicaceae	*Arabidopsis lyrata*	6 (a-f)
		Arabidopsis thaliana	5 (a-e)
		Brassica napus	4 (a-d)
		Brassica oleracea	2 (a-b)
		Brassica rapa	2 (a-b)
	Asteraceae	*Cynara cardunculus*	1
	Caricaceae	*Carica papaya*	2 (a-b)
	Cucurbitaceae	*Cucumis melo*	6 (a-f)
	Euphorbiaceae	*Hevea brasiliensis*	1
		Manihot esculenta	6 (a-f)
		Ricinus communis	1
	Fabaceae	*Acacia auriculiformis*	1
		Glycine max	12 (a-l)
		Lotus japonicus	3 (a-c)
		Medicago truncatula	4 (a-d)
		Vigna unguiculata	1
	Lamiales	*Digitalis purpurea*	2 (a-b)
		Salvia sclarea	1
	Linaceae	*Linum usitatissimum*	10 (a-j)
	Malvaceae	*Gossypium hirsutum*	1
		Theobroma cacao	5 (a-e)
	Ranunculaceae	*Aquilegia caerulea*	2 (a-b)
	Rosaceae	*Malus domestica*	15 (a-o)
		Prunus persica	4 (a-d)
	Rutaceae	*Citrus sinensis*	3 (a-c)
	Salicaceae	*Populus trichocarpa*	9 (a-i)
	Solanaceae	*Nicotiana tabacum*	9 (a-i)
		Solanum lycopersicum	2 (a-b)
		Solanum tuberosum	5(a-e)
	Vitaceae	*Vitis vinifera*	4 (a-d)
Monocotyledons		*Aegilops tauschii*	1
		Brachypodium distachyon	3 (a, b, d)
		Elaeis guineensis	5 (a-e)
		Oryza sativa	4 (a-d)
		Sorghum bicolor	6 (a-f)
		Zea mays	5 (a-e)

The sequences were downloaded from http://www.mirbase.org/

Supplemental Table 2. Sequence of gma-miR172 family members.

Family member	Mature sequence
gma-miR172a	AGAAUCUUGAUGAUGCUGCAU
gma-miR172b	AGAAUCUUGAUGAUGCUGCAU
gma-miR172c	GGAAUCUUGAUGAUGCUGCAG
gma-miR172d	GGAAUCUUGAUGAUGCUGCAGCAG
gma-miR172e	GGAAUCUUGAUGAUGCUGCAGCAG
gma-miR172f	AGAAUCUUGAUGAUGCUGCA
gma-miR172g	GCAGCACCAUCAAGAUUCAC
gma-miR172h	GCAGCAGCAUCAAGAUUCACA
gma-miR172i	GCAGCAGCAUCAAGAUUCACA
gma-miR172j	GCAGCAGCAUCAAGAUUCACA
gma-miR172k	UGAAUCUUGAUGAUGCUGCAU
gma-miR172l	GGAAUCUUGAUGAUGCUGCAU

The sequences were downloaded from http://www.mirbase.org/.

Supplemental Table 3. Results of miR172c target prediction in psRNATarget.

Gene	Domain	Multiplicity	Inhibition	Cleavage site
Glyma01g39520	AP2	1	Cleavage	CDS
Glyma02g09600	AP2	1	Cleavage	CDS
Glyma03g33470	AP2	1	Cleavage	CDS
Glyma05g18041	AP2	1	Cleavage	CDS
Glyma11g15650	AP2	1	Cleavage	UTR
Glyma11g05720	AP2	1	Cleavage	CDS
Glyma13g40470	AP2	1	Cleavage	CDS
Glyma15g04930	AP2	1	Cleavage	CDS
Glyma12g07800	AP2	1	Cleavage	CDS
Glyma17g18640	AP2	1	Cleavage	CDS
Glyma19g36200	AP2	1	Cleavage	CDS

Note：psRNATarget : *A Plant Small RNA Target Analysis Server* *http://plantgrn.noble.org/psRNATarget/* . Multiplicity indicates the the number of miRNA/target site pairs.

Supplemental Table 4. Primers used in this study.

35S:miR172c-Forward	GGTACCCGGGGATCCCACTCATTCTACCTATTATTCCG
35S:miR172c-Reverse	GTTGGATCCCTTTAGTTTATTTAGGACTTCATTAGGTC
STTM172c-48-Forward	CGACCGGTCTGCAGCATCAC
STTM172c-48-Reverse	CGGGATCCGGAATCTTGAT
35S:GmNNC1-GFP-Forward	CCAAGCTTATGTTAGATCTTAATCTCAATG
35S:GmNNC1-GFP-Reverse	CGGGATCCCTATGGTGGTGCCTGCGG
RNAi-GmNNC1-Forward	GGGGTACCACTAGTGTCCAAGGAGGAATTCGTAC
RNAi-GmNNC1-Reverse	CGGGATCCGAGCTCGGATCAATGCCGATTCTCTC
GmNNC1-503-Forward	GCGTCGACATGTTAGATCTTAATCTCAATG
GmNNC1-503-Reverse	CGGAATTCCTATGGTGGTGCCTGCGG
MBP-GmNNC1-Forward	CGGGATCCATGTTAGATCTTAATCTCAATG
MBP-GmNNC1-Reverse	GCTCTAGACTATGGTGGTGCCTGCGG
promiR172c:GUS-Forward	GGGGATCCCACTCATTCTACCTATTATTCCG
promiR172c:GUS-Reverse	GTTGGATCCCTTTAGTTTATTTAGGACTTCATTAGGTC
proGmNNC1:GUS-Forward	CGGAATTCGATCTACTTAAGTCAATAAC
proGmNNC1:GUS-Reverse	GAAGATCTCGTGGCGCTTCTTTTTTTTA
35S:GmNNC1-Forward	CGACCGGTATGTTAGATCTTAATCTCAATG
35S:GmNNC1-Reverse	GGGGATCCCTATGGTGGTGCCTGCGG
proENOD40-1-GFP-Forward	ACTAGGTTAATTAAATGGGTTTGT
proENOD40-1-GFP-Reverse	GCTTCTTCAAGAACCATGGATGGAT
proENOD40-2-GFP-Forward	ACAGTACAGTATGGCTACATACAC
proENOD40-2-GFP-Reverse	GGATGGTTGTGAGCCAACAAAGCT
proENOD11-GFP-Forward	AGGGTAACAAAGACGACGACGATT
proENOD11-GFP-Reverse	GTTTTAGTTTCTGATGAATAACAC
qRT-NNC1-Forward	CAATGGGCAGGAAAGAGC
qRT-NNC1-Reverse	ATGGCAGTCGATGGAAAGGT
qRT-ENOD40-1-Forward	TCTCTCTTGAGTGGCAGAAGCA

qRT-ENOD40-1-Reverse	TGGAGTCCATTGCCTTTTCG
qRT-ENOD40-2-Forward	TGGAGAGAAAGGGCAGATAC
qRT-ENOD40-2-Reverse	ATTGCCTACCTACTCATCTG
qRT-NIN-Forward	TGGCGCACCATGCTAACAT
qRT-NIN-Reverse	GGGTGTCATGGCAATCCTTT
qRT-ELF1B-Forward	GTTGAAAAGCCAGGGGACA
qRT-ELF1B-Reverse	TCTTACCCCTTGAGCGTGG
qRT-CYP2-Forward	CGGGACCAGTGTGCTTCTTCA
qRT-CYP2-Reverse	CCCCTCCACTACAAAGGCTCG
qRT-miR172c-STTM-Forward	CGACCGGTCTGCAGCATCAC
qRT-miR172c-STTM-Reverse	CGGGATCCGGAATCTTGAT
SL-gma-miR172a/b/k/l-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCAC TGGATACGACATGCAG
SL-gma-miR172c-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCAC TGGATACGACCTGCAG
SL-gma-miR172d/e-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCAC TGGATACGACCTGCTG
SL-gma-miR172f-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCAC TGGATACGACTGCAGC
SL-gma-miR172g-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCAC TGGATACGACGTGAAT
SL-gma-miR172h/i/j-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCAC TGGATACGACTGTGAA
SL-miR1520d-RT	GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCAC TGGATACGACTTGTCA
gma-miR172a/b-qPCR-Forward	GGACCAGAATCTTGATGATG
gma-miR172c-qPCR-Forward	GGACCGGAATCTTGATGATG
gma-miR172d/e-qPCR-Forward	GGACCGGAATCTTGATGATGCTG
gma-miR172f-qPCR-Forward	GGACCAGAATCTTGATGAT
gma-miR172g-qPCR-Forward	GGACCGCAGCACCATCAAG
gma-miR172h/i/j-qPCR-Forward	GGACCGCAGCAGCATCAAGA
gma-miR172k-qPCR-Forward	GGACCTGAATCTTGATGATG
gma-miR172l-qPCR-Forward	GGACCGGAATCTTGATGATG
gma-miR1520d-qPCR-Forward	CGGACCATCAGAACATGACACG

gma-miRNA-universal-Reverse	GTGCAGGGTCCGAGGT
RACE-glyma03g33470-in	CGGAGTTAGTGATTGATGGTGGA
RACE-glyma03g33470-out	CCGAGGCTGAATGTTGGATACCA
RACE-glyma05g18170-in	CACCTCAATAGTAAAGTTGTGGC
RACE-glyma05g18170-out	CCGTTTGTTAGTTTGACTGACCC
RACE-glyma11g15650-in	GTGGGGGCTTAACCTAGTAATAATATTC
RACE-glyma11g15650-out	AAGCAAGAGACAGCAATAAGGTG
RACE-glyma11g05720-in	TCCCTTCTTGTCTTTCACCC
RACE-glyma11g05720-out	TCCAGAGCGAGTTCCCAAAT
RACE-glyma12g07800-in	TTCAGATACATTGCGGGTAGGCG
RACE-glyma12g07800-out	CAGAAACATCGTTGTGCCAGTGC
RACE-glyma13g40470-in	GCGTTGCTACCTGGTGTGCTGGG
RACE-glyma13g40470-out	CCTAAGAGAACTGGAGCACTCAA
RACE-glyma15g04930-in	TGGGATGGTGCCTGCGAGGACTT
RACE-glyma15g04930-out	CTTCAGAACTTGTGTGGTGGGCT
RACE-glyma17g18640-in	GGGTGGAACCGTTGAAGGGTAAT
RACE-glyma17g18640-out	ATAGTAAAGTTGTGGTCTTGTGTGAGC
RACE-glyma19g36200-out	GCAGAGTTCAAGAAGCCATCGGG
RACE-glyma01g39520-out	TGGCTCTCCCTATTCTTTCACCCA
RACE-glyma02g09600-out	CGGGTAGCCATTGTAAGCAA
RT-GmNNC1-Forward	ATGTTAGATCTTAATCTCAATG
RT-GmNNC1-Reverse	CGGGATCC CTATGGTGGT GCCTGCGG
Bar-Forward	AAGGATAGTGGGATTGTGCG
Bar-Reverse	AGTCGGGAAACCTGTCGTG
G4DBD-Forward	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAA GCTACTGTCTTCTATCG
G4DBD-Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCTATTG ATTCGACCTCGACGATAC
GmNNC1-503-Forward	GCGTCGACATGTTAGATCTTAATCTCAATG
GmNNC1-503-Reverse	CGGAATTCCTATGGTGGTGCCTGCGG
GmNNC1DBD-Reverse	GGGGACCACTTTGTACAAGAAAGCTGGGTCTATG GTGGTGCCTGCGG

ENOD40-ChIP-1-Forward	AAGTAATAAGCAAATGGATAAT
ENOD40-ChIP-1-Reverse	GCTTCTTCAAGAACCATGGATGGAT
ENOD40-ChIP-2-Forward	TTGCTGGCATGGCTGGTGAGAG
ENOD40-ChIP-2-Reverse	ACTCAAAGCCCACTTTGCTTTC
ENOD40-ChIP-3-Forward	ATCAAAAGTCAAGCACGAATG
ENOD40-ChIP-3-Reverse	TCCTCGCCAAATTTCAGCTT
ENOD40-ChIP-4-Forward	AGTGTGTAAGATACAATTAGGT
ENOD40-ChIP-4-Reverse	ATTATCCATTTGCTTATTACTT
ENOD40-ChIP-5-Forward	ACAGTATGGCTACATACAC
ENOD40-ChIP-5-Reverse	AATATGTTACGAGGATTTGT
ENOD40-ChIP-6-Forward	ATTAGTGGATAGTAAGGAATTT
ENOD40-ChIP-6-Reverse	ATAATCCTCTTAAGTATTGAAT
ENOD40-ChIP-7-Forward	CTTGGAGGCTAAAACATATG
ENOD40-ChIP-7-Reverse	AGGCTTTTATTGCAGCAGTG
ENOD40-ChIP-8-Forward	ATTCCCATGTAAACCAGTAGAAG
ENOD40-ChIP-8-Reverse	GAAAGAAAATTGTACTGCGCTC
probe-ENOD40-1-Forward	CTTTTAGCAAATCCTCGTAACATATTTCTA
probe-ENOD40-1-Reverse	TAGAAATATGTTACGAGGATTTGCTAAAAG
probe-ENOD40-2-Forward	CGGTAATTAGGTTAAGGTTTAGATTCAATA
probe-ENOD40-2-Reverse	TATTGAATCTAAACCTTAACCTAATTACCG