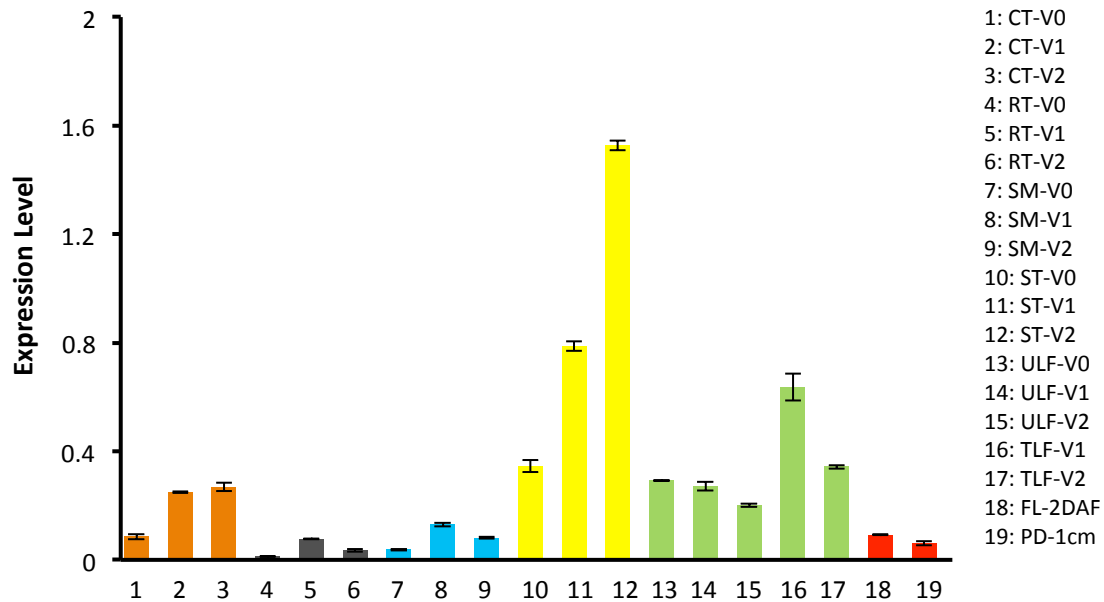
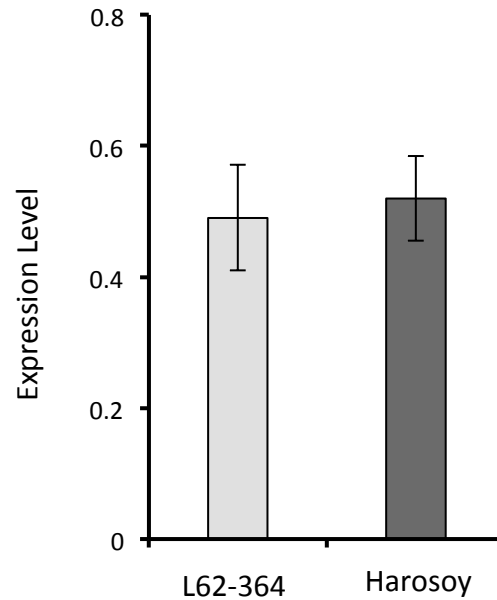


AGL79	MGRGRVQLRR	ENK	IRRQVTFS	KRRRT	TGLVKKAAQE	ISVL	CDAEVAL	IVFSP	HKG	LFEYS	AGSSME		
CAL	MGRGRVQLKR	ENK	INRQVTFS	KRRRT	GLLKKAQE	ISVL	CDAEVAL	IVFS	HKG	LFEYS	ES	CME	
AP1	MGRGRVQLKR	ENK	INRQVTFS	KRRRA	GLLKKAHE	ISVL	CDAEVAL	VVFS	HKG	LFEYS	TD	CME	
FUL	MGRGRVQLKR	ENK	INRQVTFS	KRRRS	GLLKKAHE	ISVL	CDAEVAL	IVFS	SKG	LFEYS	TD	CME	
Glyma18g50910	MGRGRVQLKR	ENK	TSQQVTFF	KRRS	GLLKKASE	ISVL	CDAQVAL	IIFST	KG	LFEYS	ERS	ME	
Glyma01g08150	MGRGRVQLKR	ENK	INRQVTFS	KRRRS	GLLKKAHE	ISVL	CDAEVAL	IVFS	HKG	LFEYAT	DS	CME	
Glyma02g13420	MGRGRVQLKR	ENK	INRQVTFS	KRRG	GLLKKAHE	ISVL	CDAEVAL	IIFSH	KG	LFEYAT	DS	CME	
Glyma05g07380	MGRGRVQLKR	ENK	INRQVTFS	KRRRS	GLLKKARE	ISVL	CDADVAL	IVFS	TKG	L	DYS	NQP	CTE
Glyma06g22650	MGRGRVQLKR	ENK	INRQVTFS	KRRS	GLLKKAHE	ISVL	CDAEVAL	IVFS	TKG	LFEYS	SD	PCME	
Glyma08g27680	MGRGRVQLKR	ENK	TSQQVTFS	KRRS	GLLKKAHE	ISVL	CDAQVAL	IMFST	KG	LFEYS	ERS	ME	
Glyma08g36380	MGRGRVQLKR	ENK	INRQVTFS	KRRRA	GLLKKAHE	ISVL	CDAEVAL	IVFS	HKG	LFEYAT	DS	CME	
Glyma16g13070	MGRGRVQLKR	ENK	INRQVTFS	KRRRA	GLLKKAHE	ISVL	CDAEVAL	IVFS	HKG	LFEYAT	DS	CME	
Glyma17g08890	MGRGRVQLKR	ENK	INRQVTFS	KRRS	GLLKKARE	ISVL	CDADVAL	IVFS	TKG	L	DYS	NE	PCME

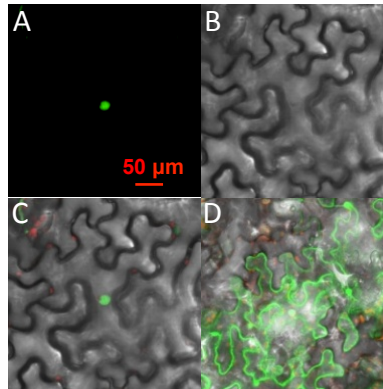
**Supplemental Figure 1.** Alignment of predicted MADS-box domains of the *Dt2* candidate gene homologs in soybean and *Arabidopsis*.



**Supplemental Figure 2.** Expression of the *Dt2* candidate genes in the semi-determinate soybean cultivar NE3001 detected by qRT-PCR. The y-axis indicates expression of *Dt2* relative to expression of *Cons4* in different tissues including cotyledon (CT), roots (RT), stems (SM), stem tips (ST), unifoliate leaflets (ULF), trifoliate leaflets (TLF), flowers (FL), and 1-cm immature pod (PD) at different developmental stages including V0 (when the cotyledons at node 0 are fully extended but the unifoliate leaflets at node 1 are not yet unrolled), V1 (unifoliate leaflets at node 1 fully expanded, but 1st trifoliate leaflets at node 2 not yet unrolled), and V2 (the first trifoliate leaflets have fully unrolled but 2<sup>nd</sup> trifoliate leaflets are still unrolled) stages as shown in x-axis. Expression levels were shown as means  $\pm$  standard errors of the means from four replicates.

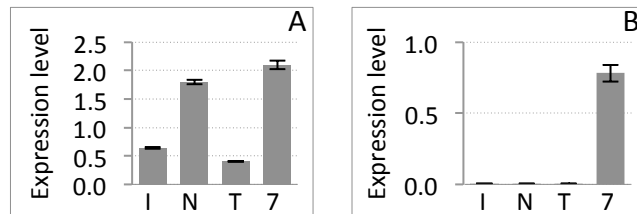


**Supplemental Figure 3.** Expression of *Glyma18g51000* in apical stem tips of NILs L62-364 and Harosoy at V2 stage detected by qRT-PCR. The y-axis indicates the expression levels of the gene relative to expression of *Cos4*. Expression levels were shown as means  $\pm$  standard errors of the means from four replicates.



**Supplemental Figure 4.** Subcellular localization of the Dt2 protein in tobacco epidermal cells.

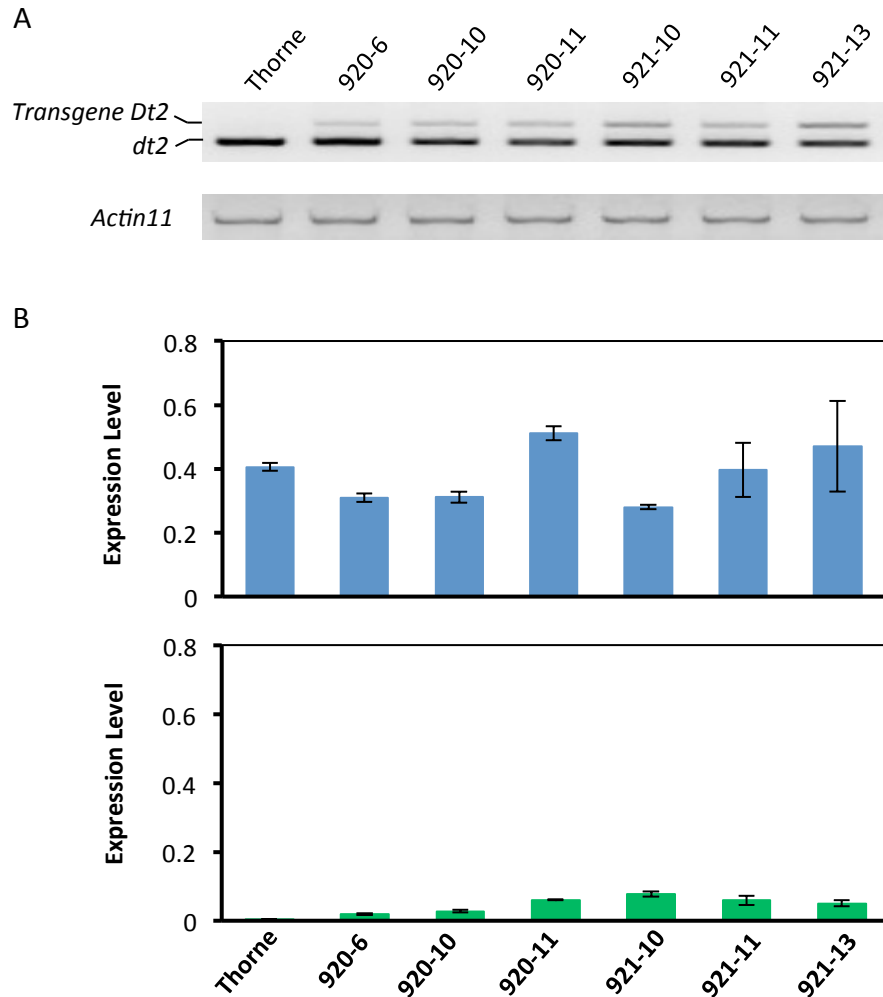
- (A) Subcellular localization of *Dt2*-GFP fusion gene under the control of 35S promoter as observed with a dark field for green fluorescence.
- (B) The same cell shown in (A) as observed with a bright field for the cell morphology.
- (C) The merged image of (A) and (B).
- (D) Subcellular localization of GFP protein as a control.



**Supplemental Figure 5.** Expression of endogenous *Dt2/dt2* and/or the transgene *Dt2* in parental and transgenic lines determined by qRT-PCR.

**(A)** Expression of *Dt2/dt2* relative to expression of *Cons4* in apical stem tips of IA3023 (I), NE3001 (N), Thorne (T), and a T3 Thorne *Dt2* transgenic plant as shown in Figure 3 at the V2 stage determined by qRT-PCR in . Expression levels were shown as means  $\pm$  standard errors of the means from four replicates.

**(B)** Expression of transgene *Dt2* relative to expression of *Cons4* in the same samples as shown in (A).



**Supplemental Figure 6.** Expression of Thorne endogenous *dt2* and the transgenic *Dt2* in the Pro-*Dt2*:CDS-*Dt2* transgenic lines determined by qRT-PCR. A. RT-PCR products of Thorne native *dt2* and transgenic *Dt2* amplified from Thorne and all six transgenic lines with a pair of *Dt2/dt2* specific primers. The small fragments were PCR products amplified from the native *dt2*, whereas the larger fragments were amplified from the transgene *Dt2* in six transgenic lines from distinct transformation events. Gene *Actin11* was used as a control. B. Expression of Thorne *dt2* and transgene *Dt2* relative to expression of *Cons4* detected by qRT-PCR with a pair of *Dt2/dt2* specific primers (top plot), and specific expression of the transgene *Dt2* relative to expression of *Cons4* detected by qRT-PCR with one primer from the coding sequence of *Dt2* and the other from pPTN1178 cassette.

**Supplemental Table 1.** Molecular markers used for mapping of the *Dt2* gene.

Marker	Chromosome ID	Forward primer sequences (5' to 3')	Reverse primer sequences (5' to 3')	Type of markers
SSR_18_1791	Gm18	TGACCAGTCAATTGTTCAATTCTTT	TTTACTCAACCATCTCCGCA	SSR
SSR_18_1807	Gm18	TCATTCTGTAAAATGAGTTGTGTATTC	TTATTTTGCTTTCAAACCTTACAATTC	SSR
SSR_18_1817	Gm18	GTGAGGCCATCAATCACCTT	CGCAAGAAGAAAAGAAAAGGAA	SSR
SSR_18_1821	Gm18	GGTGCCTTTAATTTCTTTGGA	ATTCACCAGATCATGTGCCA	SSR
SSR_18_1822	Gm18	AATTTGATGCACTTGATAACGA	TGACAAACACAAGAACTCACACA	SSR
SSR_18_1825	Gm18	GAATCCACCATCACCAAACC	CAATGGCAACCCAGTAAGGT	SSR
SSR_18_1831	Gm18	TGTTTTTGTTAAATCTTTTGTTTGG	TGTGTATGTTTGTGTGTGCACTT	SSR
SSR_18_1833	Gm18	GGCTATTGCAACATTCGGTT	GAGGAAAGTGTTCAATTGCCG	SSR
SSR_18_1838	Gm18	TTCTATATTCAAACCTGAACTGAACTG	AACTTATTATAACGCAATTTTATGCTT	SSR
SSR_18_1842	Gm18	TGAAATGGAGGAGAAAATGGA	GTCCGGGGAAACTGAACC	SSR
SSR_18_1846	Gm18	CTTTTAACGATTGGGTTGGG	CTTCGGCCTTAGACTTTTCG	SSR
SSR_18_1854	Gm18	GCCACCTCTACACCAACACA	TGACCAACAATGGCTTTCAA	SSR
SSR_18_1858	Gm18	TAGCTTTATAATGAGTGTGATAGAT	GTATGCAAGGGATTAATTAAG	SSR
SSR_18_1864	Gm18	TGAATGATATATGTTTTGCGAAGA	CAATAGAGCCGGATGGATGT	SSR

SSR_18_1890	Gm18	TGTTAGTGTACGCGTTACAAAATATAA	AAAGTGCATGTACATTAGTGAATTTTA	SSR
SSR_18_1926	Gm18	TTTGAGATTACTGACAAAAGAGA	TTTTGTCCCTTAAAATAACTTCAAC	SSR
SNP1	Gm18	CTCTGTAATATGCTCAGAGTC	GTAGGTGGCAAGAAACCCCCC	SNP
SNP2	Gm18	CAGACATAATCTATGAACAAG	GCAAACAACCTAAAGGATCACAG	SNP
SNP3	Gm18	CCATGTACATTAGTATTCAGTAG	AGCAGCTCTGAAATTAGCC	SNP
SNP4	Gm18	GTGTTTATATTAGTTCTTTACCC	ACCATGTATAAATGATAC	SNP
SNP5	Gm18	CAAGCACTATAGCCTTTAGTC	AGAAGCATTCTTTGAAGAGGAAAC	SNP
SNP6	Gm18	TGAAGCGGATCGAGAACAAAACA	AATGATGAACGAGTAGGAACCT	CAPS

---



**Supplemental Table 2.** Genes in the defined *Dt2* region according to the Williams 82 reference genome.

<b>Genes</b>	<b>Annotation</b>
<i>Glyma18g50910</i>	MADS box transcription factor
<i>Glyma18g50920</i>	Uncharacterized conserved protein, contains RCC1 domain
<i>Glyma18g50930</i>	MEKK and related serine/threonine protein kinases
<i>Glyma18g50940</i>	DSBA-like thioredoxin domain
<i>Glyma18g50950</i>	Ring finger protein 24-related
<i>Glyma18g50960</i>	No functional annotations
<i>Glyma18g50970</i>	Pollen proteins Ole e I like
<i>Glyma18g50980</i>	Glucosyl Transferases
<i>Glyma18g50990</i>	F-box domain
<i>Glyma18g51000</i>	F-box domain

**Supplemental Table 3.** Polymorphisms of a single nucleotide variant in the coding region of the *Dt2* candidate gene in a natural population previously described<sup>a</sup>

Accessions	Genotype	Phenotype	Type	Region	Country	Maturity Group	SNP
PI483464A	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Ningxia	China	III	G
PI 407301	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Jiangsu	China	V	G
PI 483465	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Shaanxi	China	V	G
PI468400A	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Ningxia	China	IV	G
PI 407131	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Kumamoto	Japan	VI	G
PI 447004	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Jilin	China	III	G
PI 366120	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Akita	Japan	IV	G
PI 407170	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Kyonggi	Korea, South	V	G
PI 549046	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Shaanxi	China	III	G
PI 407140	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Kumamoto	Japan	VII	A
PI326582A	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Primorye	Russia	II	A
PI 464935	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Jiangsu	China	VI	A
PI 468916	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Liaoning	China	III	A
PI339871A	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Cheju	Korea	V	A
PI 458538	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Heilongjiang	China	OOO	A
PI597459D	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Shandong	China	III	A
PI 393551	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Taiwan	China	X	A
PI597461A	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Shandong	China	IV	A

PI 562559	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Cholla Puk	Korea, South	V	A
PI 407282	<i>Dt1</i>	indeterminate	<i>Glycine soja</i>	Cheju	Korea, South	VI	A
PI 548603	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Indiana	USA	IV	G
PI 548488	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Missouri	USA	V	G
PI 548311	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Ontario	Canada	O	G
PI 548379	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Heilongjiang	China	O	G
PI 548298	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Unknown	China	III	G
PI 548318	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Jilin	China	III	G
PI 548348	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Unknown	China	III	G
PI 548362	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Unknown	Unknown	III	G
PI 548391	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Liaoning	China	II	A
FC 33243	<i>Dt1</i>	indeterminate	N. Am. Ancestor	Unknown	Unknown	IV	A
PI 548406	<i>Dt1</i>	semideterminate	N. Am. Ancestor	Jilin	China	II	G
PI 548382	<i>dt1</i>	semideterminate	N. Am. Ancestor	Liaoning	China	OO	G
PI 548485	<i>dt1</i>	determinate	N. Am. Ancestor	Jiangsu	China	VII	A
PI 548477	<i>dt1</i>	determinate	N. Am. Ancestor	Tennessee	USA	VI	A
PI 548657	<i>dt1</i>	determinate	N. Am. Ancestor	North Carolina	USA	VII	G
PI 548445	<i>dt1</i>	determinate	N. Am. Ancestor	Jiangsu	China	VII	G
PI 548456	<i>dt1</i>	determinate	N. Am. Ancestor	Pyongyang	Korea, North	VI	A

<sup>a</sup>Hyten DL, Song Q, Zhu Y, Choi IY, Nelson RL, Costa JM, Specht JE, Shoemaker RC, Cregan PB. (2006). Impacts of genetic bottlenecks on soybean genome diversity. Proc Natl Acad Sci USA. 103:16666-16671.

**Supplemental Table 4.** Phenotypic segregation for stem growth habit of the T3 progenies from individual T2 plants derived from nine independent transformation events

<b>No. in Fig. 3</b>	<b>Transformation event</b>	<b>No. of positive T2 plants planted in greenhouse</b>	<b>No. of semi-determinate T3 plants in the field</b>	<b>No. of indeterminate T3 plants in the field</b>
1	917-70	2	4	1
2	917-49	2	13	1
3	913-15	4	19	8
4	917-46	4	17	3
5	917-56	1	6	2
6	917-55	1	3	1
7	917-66	2	12	3
8	917-24	1	5	2
9	917-65	1	6	3

**Supplemental Table 5.** Correlation between expression level of the transgenes and phenotypic variation among transgenic plants

Phenotype	Thorne	913-15	917-49	917-46	917-24	917-70	917-56	917-55	917-65	917-66	$r^a$	$p$
Node number	21	10	13	15	15	15	16	17	20	21	-0.842 <sup>**</sup> , <sup>b</sup>	0.004 <sup>b</sup>
Expression Level	1	29.5	12.2	6.2	7.4	5.9	4.2	4.3	2	2.4	-0.815 <sup>**</sup> , <sup>c</sup>	0.007 <sup>c</sup>
Plant height (cm)	82.4	26.1	33.6	39.6	40.1	41.9	42.1	42.5	55.2	57.8		

<sup>a</sup>Pearson's correlation coefficients were calculated using the SPSS statistics package

<sup>b</sup>Correlation between node numbers and expression levels of the transgene

<sup>c</sup>Correlation between plant heights and expression levels of the transgene

<sup>\*\*</sup>Correlation is significant at the 0.01 level (2-tailed)

**Supplemental Table 6.** Genes surrounding *Dt2* in soybean and their putative orthologs in *Medicago truncatula*

Query genes in soybean <sup>a</sup>	Putative orthologs in <i>Medicago</i>	BLASTP		
		Identity (%)	Alignment length	Expect value
<i>Glyma18g50900</i>	<i>Medtr7g016600</i>	82.03	256	4.00E-118
<i>Glyma18g50910 (Dt2)</i>	<i>Medtr7g016630 (Mt-FULc)</i>	72.59	259	1.00E-97
<i>Glyma18g50920</i>	<i>Medtr7g016640</i>	88.45	476	0
<i>Glyma18g50940</i>	<i>Medtr7g016650</i>	85.38	212	7.00E-108
<i>Glyma18g50950</i>	<i>Medtr7g016840</i>	76.99	226	3.00E-102
<i>Glyma18g50960</i>	<i>Medtr7g016900</i>	82.78	790	0
<i>Glyma18g50970</i>	<i>Medtr7g016950</i>	63.46	301	1.00E-75
<i>Glyma18g51040</i>	<i>Medtr7g016960</i>	77.88	660	0
<i>Glyma18g51050</i>	<i>Medtr7g016970</i>	70.07	441	0
<i>Glyma18g51060</i>	<i>Medtr7g017100</i>	78.02	2384	0

<sup>a</sup>Protein sequences of 40 genes flanking *Dt2* (20 upstream of *Dt2* and 20 downstream of *Dt2*) in the soybean reference genome were used to search against the protein sequences of all genes annotated in the *Medicago truncatula* genome by BLASTP