Fig. S1

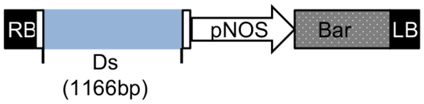

Fig. S1. The diagram of the *Ds* system.

Fig. S2

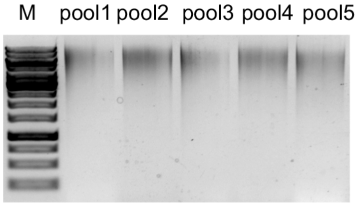

Fig. S2. Agarose gel electrophoresis of sheared DNA. 8 μg genomic DNA samples in 150 μl ddH_2O were fragmented to 8 kb

Fig. S3

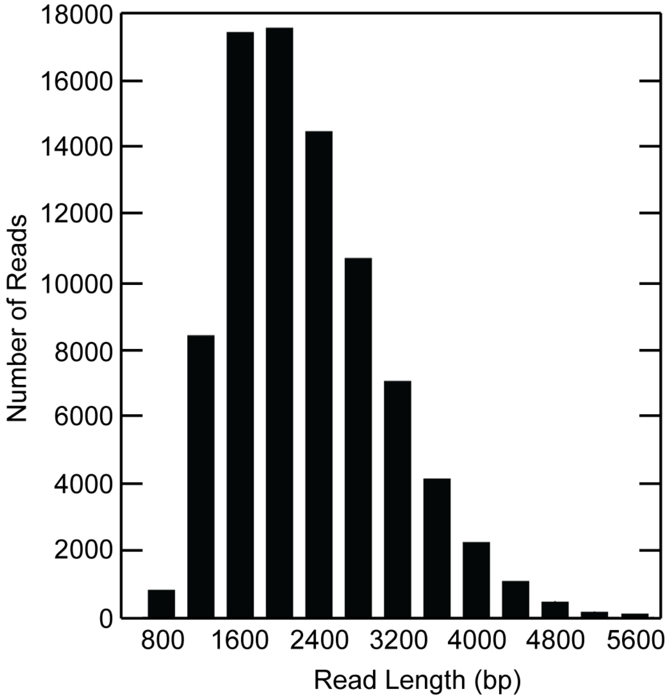

Fig. S3. Size distribution of readings containing the *Ds* elements.

Fig. S4

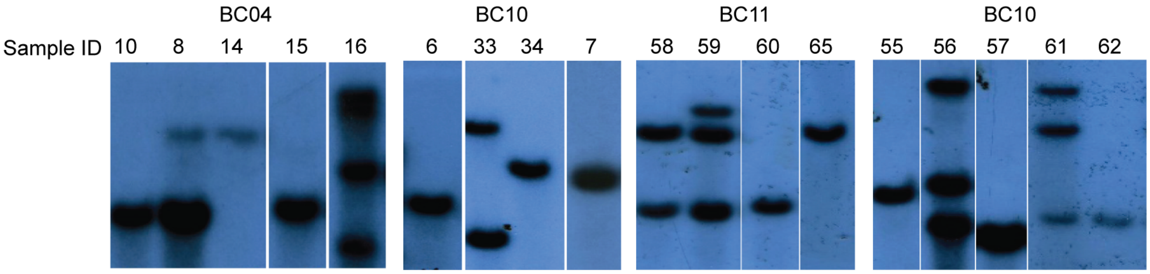

Fig. S4. Copy numbers of the transgenes in various soybean transgenic lines determined by Southern Blot. DNAs from various plants were digested with the EcoRI restriction enzyme and the T-DNA insertions were detected using probes recognizing the Cauliflower mosaic virus 35S promoter.

Table S1: DNA oligos used in this study.

Name	Sequence	Application
Primers for constructs		
BC01	AAGAAAGTTGTCGGTGTCTTTGTG	barcoding
BC02	TCGATTCCGTTTGTAGTCGTCTGT	barcoding
BC03	GAGTCTTGTGTCCCAGTTACCAGG	barcoding
BC04	TTCGGATTCTATCGTGTTTCCCTA	barcoding
BC05	CTTGTCCAGGGTTTGTGTAACCTT	barcoding
BC06	TTCTCGCAAAGGCAGAAAGTAGTC	barcoding
BC07	GTGTTACCGTGGGAATGAATCCTT	barcoding
Primers for qRT-PCR		
DSF3	TGTGTGATTTTACCGAACAAAAATAC	target region
DSR	ATAGGGATGAAAACGGTCGGTAACG	
Gm07F-HN	TAAGAGGGTCATGTACATTGTGAGAA	internal control
Gm07R-HN	AACACAGGGAATAAGACACAATCTAG	
Primers for validation		
Gm15G128500.1F2	CACAACATGTAGTAAGCCTAATTTCAA	Line 1 validation
Gm15G128500.1R2	AGTATTCTGGTTACTGATGAAGCTGTT	
Gm19G105100F	TTAGTCAAAACTCAAAAGTCTTCTCGT	Line 2 validation
Gm19G105100R	TTGAGTCTATGGAGTAAGGTGGTTAGT	
Gm11G247400F	GCAATTTGGTTTATTTATCTGACTCAC	Line 3 validation
Gm11G247400R	ATCTATGGTAAACAACAATGCCTCTAT	
Gm15G128600F	AGTATTCTGGTTACTGATGAAGCTGTT	Line 13 validation
Gm15G128600R	ATAAACTAATTTAAGGTGCATCAATGG	
Gm05G163800F	GATCACACTTTTAAGAGGAATTTGATT	Line 14 validation
Gm05G163800R	TCATCATAATCATCATCATAGAAATCG	
Gm11G181700F	CTTTTACAGGTTCTGAGGATGTAAGTC	Line 15 validation
Gm11G181700R	GGTTACATGCATACCATCTATATTTCC	
Probe sequence		
Ds-probe	Bio/ACCAGGCGAAGAAAAGAATCTTGAAGACGTAAGCACTGACGACAACAATGAAAAGAAGAAGATAAGGTCGGTGATTGTGAAAGAGACATAGAGGACACAT	target enrichment

Table S2. Positions of *Ds* insertions identified in the 15-sample sequencing

	Line number	Genomic position	Read count	P-value
DNA pool 1	Line 4	Glyma01G42560	678	0
	Line 5	Glyma17G16741	214	4.03×10^{-209}
	Line 6	Glyma02G21350	127	4.78×10^{-98}
DNA pool 2	Line 7	Glyma09G31520	5910	0
	Line 8	Glyma01G42560	8125	0
	Line 9	Glyma06G08505	11051	0
DNA pool 3	Line 10	Glyma17G09910	1110	0
	Line 11	Glyma06G08505	2965	0
	Line 12	Glyma03G31420	1285	0
DNA pool 4	Line 13	Glyma15G13491	14318	0
	Line 14	Glyma05G29675	2933	0
	Line 15	Glyma11G18750	2216	0
DNA pool 5	Line 16	Glyma01G42560	2596	0
	Line 17	Glyma19G27930	1542	0
	Line 18	Glyma03G31420	2127	0

Table S3. Positions of *Ds* insertions identified in the 51-sample sequencing

	Line number	Genomic position	Read count	P-value
DNA pool 1	Line 19	Glyma14G08990	367	0
	Line 20	Glyma01G00430	214	$8.74x10^{-164}$
	Line 21	Glyma04G41125	144	$2.04x10^{-88}$
	Line 22	Glyma20G24766	105	$1.83x10^{-52}$
	Line 23	Glyma16G029400	54	$1.77x10^{-15}$
	Line 24	Glyma11G36210	38	$1.87x10^{-7}$
	Line 25	Glyma16G02846	18	$1.83x10^{-1}$
	Line 26	Glyma15G (46980528-	18	$1.83x10^{-1}$
DNA pool 2	Line 27	Glyma03G (5158136-	23	$3.56x10^{-22}$
	Line 28	Glyma11G34210	7	$7.14x10^{-5}$
	Line 29	Glyma02G412470	7	$7.14x10^{-5}$
	Line 30	Glyma15G43170	5	$2.40x10^{-3}$
	Line 31	Glyma09G41490	4	$1.13x10^{-2}$
	Line 32	Glyma11G36210	4	$1.13x10^{-2}$
	Line 33	Glyma07G16300	4	$1.13x10^{-2}$
	Line 34	Glyma07G17370	3	$4.51x10^{-2}$
DNA pool 3	Line 35	Glyma01G03260	314	0
	Line 36	Glyma03G (5158131-	115	$1.96x10^{-91}$
	Line 37	Glyma05G36050	37	$6.38x10^{-15}$
	Line 38	Glyma16G (769719-769733)[a]	28	$4.03x10^{-9}$
	Line 39	Glyma07G17370	15	$6.21x10^{-3}$
	Line 40	Glyma04G40780	14	$1.34x10^{-2}$
	Line 41	Glyma19G02591	9	$2.53x10^{-1}$
	Line 42	Glyma13G (1615654-	9	$2.53x10^{-1}$
DNA pool 4	Line 43	Glyma09G32070	1374	0
	Line 44	Glyma16G (32671274-	852	0
	Line 45	Glyma15G18571	403	0
	Line 46	Glyma15G10791	154	$6.92x10^{-74}$
	Line 47	Glyma07G17350	107	$4.50x10^{-38}$
	Line 48	Glyma13G01110	89	$1.59x10^{-26}$
	Line 49	Glyma13G01110	70	$5.01x10^{-16}$
	Line 50	Glyma02G41957	65	$1.33x^{10-13}$
DNA pool 5	Line 51	Glyma15G18571	341	0
	Line 52	Glyma11G37670	265	$2.06x10^{-233}$
	Line 53	Glyma02G (7668714-	181	$2.94x10^{-132}$
	Line 54	Glyma17G34380	143	$1.15x10^{-91}$
	Line 55	Glyma08G44010	114	$1.44x10^{-63}$
	Line 56	Glyma03G41100	108	$3.98x10^{-58}$
	Line 57	Glyma16G (7668714-	91	$1.65x10^{-43}$
	Line 58	Glyma05G37440	73	$1.62x10^{-29}$
	Line 59	Glyma18G (32671286-	67	$2.99x10^{-25}$
DNA pool 6	Line 60	Glyma04G41100	3555	0
	Line 61	Glyma11G36210	1622	0
	Line 62	Glyma20G03910	156	$6.09x10^{-35}$
	Line 63	Glyma11G (6956901-	152	$6.46x10^{-33}$
	Line 64	Glyma11G09070	139	$1.91x10^{-26}$
	Line 65	Glyma11G07661	92	$9.96x10^{-9}$
	Line 66	Glyma11G03370	79	$2.27x10^{-5}$
	Line 67	Glyma11G (39121891-	73	$4.19x10^{-4}$
	Line 68	Glyma08G13510	71	$1.00x10^{-3}$
	Line 69	Glyma15G01120	62	$2.67x10^{-2}$

[a] means that the *Ds* insertion site localizes in the intergenic region.

Table S4. Sequencing result of T-DNA insertion lines

Pools	Samples	Southern insertion #	Identified locations	Number of reads
BC012	55 56 57 61 62	1 3 1 3 1	Chr17:6720565-6720741 Chr8:1883683_1883811 Chr16:26916636-26916774 Chr14:6470983-6471174 Chr15:9255527-9255654 Chr1:1656868-1657006 Chr17:6746122-6746317 Chr11:36293284-36293325 Chr17:10072860-10072999	189 113 102 94 38 27 13 11 10
BC011	58 59 60 65	2 3 1 1	Chr7:39843827-39844019 Chr19:1778510-1778664 Chr16:26916638-26916723 Chr11:36293284-36293442 Chr4:47915113-47915165 Chr7:39844036-39844222 Chr11:36292812-36292827	153 145 94 15 12 8 7
BC010	6 34 7 33	1 1 1 2	Chr9:38515015-38515215 Chr14:45970969-45970988 Chr9:38515223-38515420 Chr14:45970754-45970951 Chr2:35224556-35224587 Chr9:38515439-38515621	467 82 19 9 8 8
BC04	10 8 14 15 16	1 2 1 1 3	Chr19:48852620-48852751 Chr3:37108399-37108520 Chr2:588378-588482 Chr19:48851289-48851375 Chr3:37108764-37108917 Chr8:45198635-45198660 Chr17:3073403-3073433 Chr19:48851082-48851170	90 71 38 29 5 4 3 3