Supplementary Data

Figure Legends

Figure S1. The geographic distributions of landraces used in this study. Each circle represents a different province. The number of landraces per province is indicated by closed portion of the circle. The four colored portions divide China into four regions: NER (northeast region), NR (north region), HR (Huanghuai region) and SR (south region).

Figure S2. Multiple sequence alignment result of promoter sequences of $GmRCA\alpha$ and $GmRCA\beta$ by ClustalX v.1.83 program. A 2300-bp fragment upstream of the ATG start codon was designated as the promoter region.

Figure S3. The expression levels of two RCA genes in the natural population. A. The distribution of the expression levels of two RCA genes in each landrace. B. A comparison of the average value of two RCA genes in 219 soybean landraces.

Figure S4. Sequence analysis of the $GmRCA\beta$ promoter in soybean. A 2300-bp fragment upstream of the ATG start codon of $GmRCA\beta$ was designated as the promoter region. Parts of the putative cis-regulatory elements are noted under the sequences in shadow letters. The ATG is indicated in red. The nucleotide at position +1 (bold A in the gray box) is the ATG start codon. The polymorphism sites are underlined, and the cis-acting elements located in polymorphism sites are indicated in yellow.

Figure S5. Sequence analysis of the $GmRCA\alpha$ promoter in soybean. A 2300-bp fragment upstream of the ATG start codon of $GmRCA\alpha$ was designated as the promoter region. Parts of the putative cis-regulatory elements are noted under the sequences in shadow letters. The ATG is indicated in red. The nucleotide at position +1 (bold A in the gray box) is the ATG start codon.

Figure S6. Phylogenetic analysis of $GmRCA\beta$ promoter sequence. The tree was constructed with the Bayesian approaches using MrBayes v3.1.

Figure S7. PCR product and restriction analysis of 1381Z-GroupA promoter,1381Z-GroupB promoter. A. Amplified products of genomic DNA of different $GmRCA\beta$ promoter sequence in soybean. B. Vector pCAMBIA1381Z restriction results.



Figure S1. The geographic distributions of landraces used in this study. Each circle represents a different province. The number of landraces per province is indicated by closed portion of the circle. The four colored portions divide China into four regions: NER (northeast region), NR (north region), HR (Huanghuai region) and SR (south region).

GmRCA α GmRCA β	-ATTAATAGTAA-ATATTTTGTGAAATTTG TTATTTACTAAATAATTATAATAATTAAAAAGTTAATAATTATGTGTAAAAAGAATCAGTCCACTTTAACGCATCTTTGGATCTG TAGCCATAATCCTATATTACAGTTAGAGTCACCAAACCTTTTCAACTGGAT-AAAAAGTCCGTGTTGATTCTTATTTGAGTCTTGTGGGATCAQGTTAACATAAAA * *** * ***** ** ** * * * * * * * *
GmRCA a	TTCCAATCTAAAAATATGTTTGAAAATAAAATCAAATCA
GmRCA β	TGTCAAT TAT TGCATC TGA TCATTACTT TAT TCATGTGCAGGT TT TAGG TAC TTG TTT CCAAGT TAT CTG AAT TAC TCTATG TATGCATAGATGAAGG TCTAOG TCTTAT
Onucces p	
GmRCA a	
	ТТАААААТТТТТААGTСАААСАТСАГСАСТТАААСАТАТТСАТТТТААТТТТААТССААССТСАТСТААТАА
GmRCA β	TTAGAGAAATGAAATTCATAAATTCACCATATAT-GACTTAATTGTTTTAACCTTTGAAATTTAAAAATATTTTTTAGTTCTTGAATTTTGGAAAATTGCTTTTTAGTCTT
	*** * ** * ** *** * *** * * **** ** ****
GmRCA a	AAATAATGATTTTTTTATTATCTTGTCGTTAGTAAAAAAATTGATTAACTTTGTCAATAATTAATATACAAAATACTTAATCGACTATCTTTAATAAGATCTCTTAAATCATA
GmRCA β	${\tt TTGTTGATACTTTGTCACGGTTTTTATTGTCATAAATTGTTTTTTATTTTTTAGCTGCCAAAATTTATATTTTAAAACTACTAATTAGTATTAGTTTTCTGGAGTCTTTTGGCGCAGA$
	* * *** * * * * * * * * * * * * * * * *** *
GmRCA a	AAACTCAGATAGTGAGACCGAGTCCGAATACATTATCACTAGAACTAAGGATATGCTGGTGTTGAA-AATAATGAATTTGATAAACTAGAA-GTAAATTAACAATATTAAA
GmRCA β	AGAAATGAGGTCTGGGAAGGTATTCAAAAACCTCCAGAACTTTCTCTTCAATTATCAAAACTOGCACTCTCTGAATGGCAGCAACTTCAATCAATTACTCGCCGCCCTTATATTTATT
	* * ** ** * * ** ** ** ** ** ** ** ** *
GmRCA a	ATTTAATTTCACAAGCTAATTCCTAAAATATTATTATATTATAATTAAAGATGATAAAATAAAATATAGTTTAGAATATTAT
GmRCA β	A TGT GTA TGC TT AGC TCC AAAAAAAAAA TACT-G TAT TTG CTAGTAAAG ATGT GAAGG AAT TTC CAG ATT TTT CCT TAGC CCC TATGCATAACTAGT ATT TTTAGTGC TGG
onatori I.	** * * * **** **** *** ** ** ** ** * * *
GmRCA a	ATOGGTTAGTTTATTCCAAAAATAGATTGAAAAGGTCAACTTAGCTTCTTTTTTTT
GmRCA B	ACCITITATETITATETIC FORMAGCATAACATECITATEGATGGATCAATTETITATETIAACCITOGACTATITETITTETIATTIGATAATTAACAAGATGTGTGTGGCGACTAT
GILICA P	
GmRCA a	AAATGTTTAAAATAAAGGAATATGCATOGTOGAGTGAAOGTATGAGACCACTCATATTTTTGAATTTTTTTCTCTCGCCTATTTTTTTACAATCAACTCACCTTTAGATAAG-A
GmRCA β	GATTGTCTGCGGTCATAATCCAAGAACTTCATATATTTCATTAATTTGTGATGTTTGGAAAATAAAACATAGGGCAGAGATTCAACTTTCCCTCTCTCT
	* *** * * * * * * * * * * * * * * * * *
GmRCA a	${\tt TAGTCTTATGCTCA-aTTTAAAAG-ACCAAAACTTATGAAAAAAAATGCTTTGGTGAAAAATGGCCGTTAGTAAAATAAAATGAATTGAAAATTAAGAATTCAAATTTTTGAGGAAAAAAAA$
GmRCA β	CATG TGTACCTTAAGA TATACACGTGTCAAGCCATAAGTAAGATATTTCTTTATTATGAAAAGGACTAAAAAAAA
	* ** * ** ** * * * ** * *** * ** * * * *
GmRCA a	GAATCACCGGGGCACCAGCCTCGGTCTGTCGGTTGAGATAGAGCATGTTGCGAGACCTATGAATTGGATATTTTTTAATTCACCTTTTTTTT
GmRCA β	CGATCAGATAGCCAAATTOGTTCTGAGTCAGAGAGATTAATATATTCCTCCAAAAAAAACTTTTGACTAGTTTGGCAAAATGAAGGAAAACAAOGACAGAGAGAGAATTTGG
	**** * *** * * **** * **** * * *** * ** ** ** ** ** ** ** ** ** **
GmRCA a	TCTACATGAATTTATTATCACTTATTTAAGAAATTAATAACCCTCGACTTTCTTAGTAAGAGTTCATTTGAATTCACTAGTGTTTTAAAAATAAACTCTTACTTTTTTTT
GmRCA β	TTTATTTAAATATATAAAGTTTTTGGAAGCAAAAATGAAAATTTTAAACTTGCTTA-TACGGATGTATAATTAAGTAGCAATCTGAACTGATGTAATAATCTCTCTCT
	* ** * *** *** * ** *** * *** * **** * *
GmRCA a	AATATATTCTAAACTGCAAATACAAAGGCTAGTTAGGGATTAAATAAA
GmRCA B	TTTTTGAATCTTTGG TAATTCTGAAGCTTAAC-ACTGAAGAAACTGGTGAA-GTTGGGAATTTCCTTTAATTGCAGCATGAGGAATGT-TTGTGTTCCTAGAACCACATGCAACAG
onnorth	
GmRCA a	AGCAGGTAATTTGGAAGGACTAGCTCTGGAACCTTCAATTATAGCATGGCATGCTTGGAGCACATAAAGGATGTATCACTTAACCAATAAGACAGGAG-AGGCACATATAAAAGCTGA
GmRCA B	ACARGACAAGGACAAGGACAC - GACAGAGGACAT ICAA III AUCAI COUCHAI ICAI ICAI ICAI ICAI ICAI ICAI ICAI
GIRCAP	
GmRCA a	
	AAAATCACCTTGTGCGCATCCTTTGGCCCCCCCTTGGACCOGGGGGGGGGG
GmRCA β	TTAT-CCGACTCACCAGAATTAATTTATGTTTCTATTAGACAATAACTGAGAAGAAGAAGAATATGT-ACCTAAAGTGTAAAATAATTTTATACAATAAATTATTATTTAT
	* * ** * ** ** ** ** * * *** * **** * ****
GmRCA a	ATAATTTCATAATTTTTTTCAACATCATTGGCTACTTGGCTGGGTTGGCGTTAGCA-TGCTAAAAATCTAAAATAA-ATGGCAGCCCACAGATTTTTATGTAATTGACCAATACAAATAAT
GmRCA β	AAATTTTTATAAAATTATGTTAAAGGTCAAATACATTTTTGTGATTAGATAATAATAATAAAATAAAATTATTA
	* * *** *** ** * * * * * * * * * * * * *
GmRCA a	AGAAAAAAATTATGCAGTAAAAAATTTTACATTTCATTT
GmRCA β	-GAATTTTTCTCCGCCTGAAAAATACTCCAAGACTCTAAGTTAACATAATTTCATGAAAACTGAAAATTGGATCTGTTGGCTACTATGTGGGGCCAATTAGCACCATCTTTACGTAAGTAA
	*** * * * *** * * * * * * * * * * *** *** *** ** ** ** ** ** ** ** ** ** ** ** **
GmRCA a	CATA ATGTAAAAGTGTGTAACAATTAAATTCTAGAAATATTA-TAGTGTTGTTGCTGAACAATTTTCTCTTCTAGTGAAAGCAACACTAGCCAATGTTTGATCCTGATGAGCAATGTTTGATCCTGATGAGCAACACTAGCCAATGTTTGATCCTGATGAGCAACACTAGCCAATGTTTGATCCTGATGAGCAACACTTGATCCTGACTAGCCAATGTTTGATCCTGACTAGCCAATGTTTGATCCTGACTAGCCAATGTTTGATCCTGACTAGCAACACTAGCCAATGTTTGATCCTGACTAGCAACACTAGCCAATGTTTGATCCTGACTAGCAACACTAGCCAATGTTTGATCCTGACTAGCAACACTAGCCAATGTTTGATCTGATGTGACAATGTTTGACTGAC
GmRCA β	AACACCACCTCAAATATTATOGTAGTTGTTGGTTAAAAATACTAATAATATTATCATTGGTTATCGAGATTGAAGAGCACTGAAATGACCAATAACCGATGCTTGAATAAACGAC
	* * * * *** * * * ** ** ** ** ** ** * ** *
GmRCA a	ATGAATG CTACAAGCT ACG TGTGGT AACAAATTC TOCTTG AAACTT ATGGAT CAG AAAAAGGCT T CAATT CCAACCT CAAGATCT ATGGATGTAGTAGGT TGG TGAGACCTGGAATGCAC
GmRCA B	ACAAGTCCTCATAGGCTTGTGTGGCAATAAGATCTCCTTTAATACTGGTGGAGATAGAAAAAGCTTGAATACCAACGCAAAGTG-AAGGAC-TAAGACTGACAAAAAACCTTGCAGGCAC
GmRCA a	TATG TAG TIG CCACGTGGCAGCCAAGCTAAT CTT CATGTGTGGGATGTT AAGGATTCT ACAACTAGT TTTGAGGATATGAAT TT CTGAACACTCAAT-TGCACCACTGCATAAG
GmRCA B	THE INF TREAST DECLARACE ARE CLARED AND THE CONTROL AND A THE ADDALED AND THE
ouron b	
CmRCA a	
GmRCA a	
GmRCA β	TCACAAGCCACTTAGCATTGAATTTGAATTTGAAGTTAAAAGCCTTAGCCTTCCCTACAACTCCAAACACAAAGACATCATCATCATCC identity=45.2%
	ቀጥ ቀ ቀቀ ቀቀ ቀቀጥ የ የቶ የት የት የትዮጵያ የ

Figure S2. Multiple sequence alignment result of promoter sequences of $GmRCA\alpha$ and $GmRCA\beta$ by ClustalX v.1.83 program. A 2300-bp fragment upstream of the ATG start codon was designated as the promoter region.

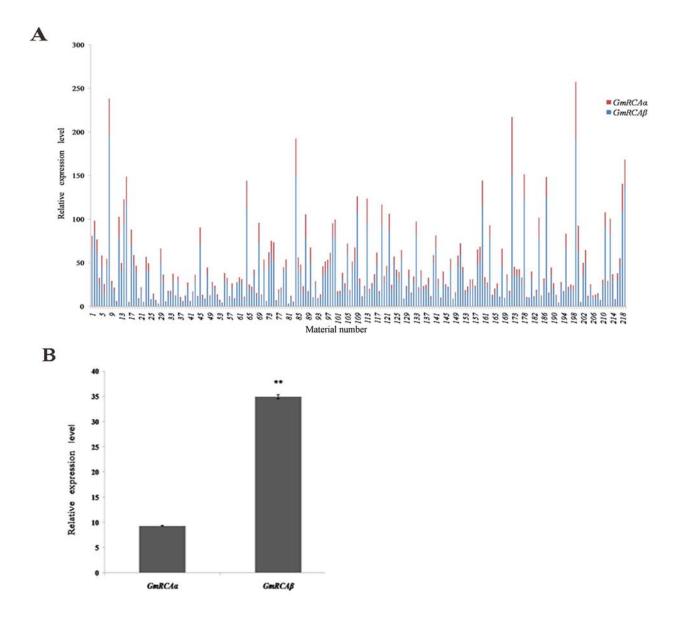


Figure S3. The expression levels of two RCA genes in the natural population. A. The distribution of the expression levels of two RCA genes in each landraces. B. A comparison of the average value of two RCA genes in 219 soybean landraces.

GmRCAB promoter region

-2300	tagccataatcctata	ttacagttagagt	cacaaaccttttca	acttggataaaaagt	ccgtgttgattctta	atttgagtcttgtgc	gatcacgttato	ca
-2200	taaatgaatattgtcaa	attattgcatctg HD-Zip	atcattactttatt	catgtgcagg <mark>ttt<u>t</u>a</mark> TATA-		agttatctgaattac	tctatgtatgca	at
-2100	agatgaaggtctacgto		atgaaattcataaa	and the second se	and the second statements	acctttgaaattta	aaaatattttt	tt
-2000	tagttettgaattttga	ataaattgtcttt	ttagtcttttgttg	atactttgtcacggt	Contraction of the second s	ttgttttttatttt 1-motif TATA-I		aa
-1900	tttatattttaaaacta	actaattagtatta	agttttctggagtc	ttttggcgcagaaga	aatgaggtetgggaa	iggtattcaaaaacc	tccagaacttto	ct
-1800	cttcaattatcaaaact	tcgcactctctga	atggcagcaacttc	aatcaattteetege	accttatatttattg	ttatgtgtatgctt	tagetecaaaaa	aa
-1700	aaaaagtactgtattts	gctagtaaagatg	tgaaggaatttcca	gatttttccttagcc	ctatgcataactagt	atttttagtgctgg	aacttttatct	tt
-1600	tcctgaaagcataacat	tgcttattgatgc	atccaatacttcta	agtttaaccgtcgac	tatttetttttttt	atttgataatttaa	ctagatgttgtg	gt
-1500	ggacttggattgtctg	cggtcataatccaa fYB Skn-1 mo		ttcattaattttgta Box 4	atgtttggaaaataa	aacatagggcagag	attcaactttco	cc
-1400	tetetetatetgeaget	taaaccacatgtg		cacgtgtcaagccat RE/G-Box GCN4-1		ttattatgaaaagg	actaaaaaata	ac
-1300	taagcaatacatttgag	gtttctttgatta	· · · · · · · · · · · · · · · · · · ·	cgatcagatagccaa	attggttctgagtca		ttcctccaaaaa	aa
-1200	aacttttgactagttt	ggcaaaatgaagg		gagatgaatttggtt GAG-motif ARI		Box 4 agtttttggaagca	aaaatgaaaatt	tt
-1100	taaacttgcttatacg	gatgtataattaa		tgatgtaataatete		ttttgaatctttcg	gtaattetgaag	gc
-1000	ttaacactgaagaaact	tggtgaagttggg	aattteetttaatt			accacatgcaacacg Box		ag
-900	gagagagacagaggaag	gaggcagagagag	gagaagaacaagga	gettgttatcategt	agggtccatggatgt	aatggggtgcaatc	accttcatgaga	at
		Internet and a second	TCA-element		AuxRR-core			
-800	tattatccgactcacce	Box 4	gtttctattagaca	MYB	gattatgtacctaaa	igtgtaaaataattt	CAAT-bo	
-700	attatttatgataaatt		taaaattatgttaa		ttgtgattagataat	aatataaaataaaa		
-600	ttttcagtgcattatca	and the second	taatccatgttgaa A-box	tttttctccgcctga	aaatactccaagact	ctaagttaacataa	tttcatgaaaad	ct
-500	gaaaattggatctgtt			atctttacgtaagta	aaacaccacctcaaa	tattatggtagttg	ttggttaaaaa	ta
	and the second s	-motif	CAAT-box	-			GT1-motif	
-400	ctaataatattatcatt	tggttatcgagat	tgaagagcactgaa		gcttgaatttgaata		tcataggettg	tg
-300	TATA-box tggcaataagatctcc Circadian	ttaatactggtgg TATA-box	agatagaaaaaagct	CAAT-box tgaataccaaccgca	aagtgaaggactaag	TGA-element agtgacaaaaaacc	ttgcaggcacag	gt
-200	atgctgccacgtgtcg		gtatgtgtggatgt	aaggettetatgace	aatccttttgtgtac	actcactctaccaa	gcaactactata	aa
					AT-box P-box		TATA-b	
	ABRE/G-B							
	gtcacagcetetggtea		cattgaatttgcaa	ttgaagttaaaagcc	ttagcettecetaca	actecaaacacaaa	gacatcatcat	cc ATG

Figure S4. Sequence analysis of the $GmRCA\beta$ promoter in soybean. A 2300-bp fragment upstream of the ATG start codon of $GmRCA\beta$ was designated as the promoter region. Parts of the putative cis-regulatory elements are noted under the sequences in shadow letters. The ATG is indicated in red. The nucleotide at position +1 (bold A in the gray box) is the ATG start codon. The polymorphism sites are underlined, and the cis-acting elements located in polymorphism sites are indicated in yellow.

GmRCA a promoter region

-2200 -+++	
-2500 attaatag Box 4	taaatatttttgtgaaatttgttatttactaaataattatatata
	${\tt tttggatctgttccaatctaaaaaatatgtttgaaaataaaatcaattcaagattatgcttaattatttgaaagactttcggacacttttgca$
2200 44080400	
-2100 aacacaat	${\tt ttacaaacataggtttaacccaaaattgagttaaaaaattttaagtcaaacatgcacttaaacatattgattttaatttaaatccaacctga$
	HSE circadian
-2000 tetaataa	aaaataatttgaatttgttttgtatttttaaaaatactaaatatagttaaaaataatgattttttattatcttgtcgttagtaaaaaaaa
-1900 gattaact	${\tt ttgtcaataattaatattcacaaaatacttaatcgactatctttaataagatctcttaaatcataaaactcagatagtgagaccgagtccga$
	Box 4
-1800 atacatta	t cactaga acta aggatatgctggtgttga aaataatga atttgata aactaga agta aatta acaatatta aa attta atttca ca agcta
-1700 attectaa	aatattattataattaaagatgataaaataaatatagtttagaatattat
	GA-motif
-1600 ggatcggt	tagtttattccaaaaatagattgaaaaggtcaacttagcttctttttttt
-1500 ataggaat	gaaataaataagaaatgtttaaattaaaaaggaatatgcatggtggagtgaaggtatgagaccactcatatttttgaatttttttt
	5UTR Py-rich stretch
-1400 gcctattt	ttttaccatcaaactcacctttagataaagatagtcttatgctcaatttaaaagaccaaaacttatgaaaaaaaa
	I-box
-1300 ccgttagt	aaaataaaatgaattgaaaattaagaattotaaatttttgagggagaatcacggggcaccagcotcggtotgtgggttgagatagagcatgt
-1200 tgcgagac	ctatgaattggatattttttaattcacctttttttttcctttcataatcagtacttagactattatctacatgaatttattt
-1100 taagaaat	taataaccctcgactttcttagtaagagttcatttgaattcactagtgttttaaaataaactcttactttttttt
Box	: 4 TCT-motif
-1000 tctaaact	gcaa a taca a a ggct a gtt a ggg att a a a ta a a tcact gct gtt a a tt g a tct a a a a ca a ga a ta a a ca a gtt a t gtt a ca a a t g g g a t a a a ca a gtt a t g t a ca a a t g g g a t a a ca a gtt a t g t a ca a a t g g g a t a a ca a g t a t g t a ca a a t g g g a t a a ca a g t a t g t a ca a a t g g g a t a a ca a g t a t g t a ca a a t g g g a t a a ca a g t a t g t a ca a a t g g g a t a a ca a g t a t g t a ca a a t g g g a t a a ca a g t a t g t a ca a a t g g g a t a a ca a g g g g g g g g g g g g g g g
	GARE-motif
-900 tcatctta	tgatagcaggtaatttggaaggactagctctggaaccttcaattatagcatgggcatgcttggagcacataaaggatgtatcacttaaccaa
Skn-1_motif	
-800 taagacag	ggagaggcacatataaaagctgaaaaaatcaccttgtgtgcatcctttggcccccctttgaaccgggagtggaggggggtcattacatcatc
700	Sp1 Skn-1_motif
-700 tcaatcaa	ccactateettgtttetattaatgeaatggeeeactttttgacataattteataatttttttteaacateattggetagttag
-600 catgetaa	Box 4
-000 catgetaa	aaatetaaataaatggeageeeageeagatttttatgtaattgaeeaataataattagaaaaaattatgeagtaaaaaattttae Box-W1 ATGCAAAT motif HSE
-500 attttcat	ttaattaactttattaacttttatgataattacttaaaaaattatattaagaacgaattatgatttacataatgtaaaagtgtgtaacaatt
and a second	chs-CMA1a TATA-box
-400 aaatteta	gaaatattatagtgttgttgctgaacaattttctcttctagtgaaagcaacactagccaatgtttgatcctgatgagatgaatgctacaagc TATA-box O2-site
	gtaacaaatteteettgaaacttatggateagaaaaaggetteaatteeaacteaagatetatggatgtagtaggtggggagaeetggaa Box CAAT-box CAAT-box
	gtagttgccacgtggcagccaagctaatcttcatcgtgtggatgttaaggattctacaactagttttgaggatatgaatttctgaacactca
	x II/ABRE/G-Box ATCT-motif/ circadian GTGGC-motif CAAT-box
	actgccataagtcccttccctattgtaatttctgccagcctcaaaaaagacttgctctgcctctgccactctctgagtaggactacactgccATG
	CAT-box +1

Figure S5. Sequence analysis of the $GmRCA\alpha$ promoter in soybean. A 2300-bp fragment upstream of the ATG start codon of $GmRCA\alpha$ was designated as the promoter region. Parts of the putative cis-regulatory elements are noted under the sequences in shadow letters. The ATG is indicated in red. The nucleotide at position +1 (bold A in the gray box) is the ATG start codon.

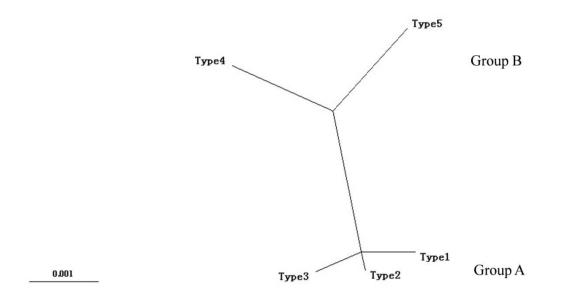


Figure S6. Phylogenetic analysis of $GmRCA\beta$ promoter sequence. The tree was constructed with the Bayesian approaches using MrBayes v3.1.

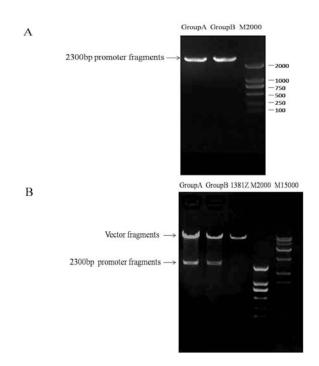


Figure S7. PCR product and restriction analysis of 1381Z-GroupA promoter, 1381Z-GroupB promoter. A. Amplified products of genomic DNA of soybean different $GmRCA\beta$ promoter sequence. B. Vector pCAMBIA1381Z restriction results.

The second se	Descriptive statistics			Variance analysis				
Traits	Mean Range		SD	Source of variation df		Sum of square	Mean of square	P value
<i>GmRCAα</i> expression	9.26	0.47-66.41		Among landraces	217	60336.78	278.05	< 0.001
			9.6	Block(time segment)	2	3.32	1.66	< 0.05
				Error	434	171.77	0.4	
				Total	653	60511.87	280.11	
				Among landraces	217	742218.73	3420.36	< 0.001
GmRCAβ	34.65	2.38-196.49	33.36	Block(time segment)	2	73.8	36.9	< 0.001
expression				Error	434	2086.55	4.81	
				Total	653	744379.08	3462.07	

Table S1. The descriptive statistics and variance analysis of GmRCAa expression and $GmRCA\beta$ expression in natural population

Table S2. Regression coefficients among $GmRCA\alpha$ expression, $GmRCA\beta$ expression and yield components in a natural soybean population

Traits	Fv/Fm ^a	$\Phi_{PSII}{}^{b}$	qP ^c	NPQ ^d	Seed weight	Seed yield
$GmRCA\beta$ expression	ns	0.0003*	0.0004**	ns	0.027*	0.025**
GmRCAa expression	ns	0.001*	0.001**	ns	ns	0.073**

^a Fv/Fm: maximum quantum yield of PSII primary photochemistry in the dark-adapted state; ^b ΦPSII: actual quantum yield in the

light-adapted state; ^c qP: photochemical quenching coefficient; ^dNPQ: non-photochemical quenching parameter describing the regulated

dissipation of excess energy; ns: not significant; *, P<0.05; **, P<0.01.

Elements type	Names of the elements	Organism	Motif sequence	Function		
	AT1-motif	Solanum tuberosum	AATTATTTTTTATT	part of a light-responsive module		
	Box 4	Petroselinum crispum	ATTAAT	part of a conserved DNA module involved in light responsiveness		
	Box II	Arabidopsis thaliana	ACACGTAGA	part of a light-responsive element		
	GAG-motif	Spinacia oleracea	AGAGATG	part of a light-responsive element		
Light	GT1-motif	Arabidopsis thaliana	GGTTAA	light-responsive element		
		Pisum sativum	CACGTT			
		Brassica napus	ACACGTGT			
	G-Box	Antirrhinum majus	CACGTA	cis-acting regulatory element involved in light responsiveness		
		Arabidopsis thaliana	CACGTG			
Circadian	circadian	Lycopersicon esculentum	CAAAGATATC	cis-acting regulatory element involved in circadian control		
	ABRE	Arabidopsis thaliana	CACGTG	cis-acting element involved in the abscisic acid responsiveness		
	AuxRR-core	Nicotiana tabacum	GGTCCAT	cis-acting regulatory element involved in auxin responsiveness		
	TGA-element	Brassica oleracea	AACGAC	auxin-responsive element		
Phytohormone	TCA-element	Brassica oleracea	GAGAAGAATA	cis-acting element involved in salicylic acid responsiveness		
	P-box	Oryza sativa	CCTTTTG	gibberellin-responsive element		
	GARE-motif	Brassica oleracea	TCTGTTG	gibberellin-responsive element		
		Glycine max	CAATT			
	CAAT-box	Hordeum vulgare	CAAT	common cis-acting element in promoter and enhancer regio		
		Arabidopsis thaliana	CCAAT			
	TATA-box	Arabidopsis thaliana	TATA			
Basal element		Brassica napus	ATATAT	core promoter element around -30 of transcription start		
Dasar element		Lycopersicon esculentum	TTTTA	core promoter element around -50 of transcription start		
	5UTR Py-rich stretch	Lycopersicon esculentum	TTTCTTCTCT	cis-acting element conferring high transcription levels		
	MBS	Zea mays	CGGTCA	MYB binding site		
Biotic stress	MID0	Arabidopsis thaliana	TAACTG	MYB binding site involved in drought-inducibility		
	ARE	Zea mays	TGGTTT	cis-acting regulatory element essential for the anaerobic induction		
	Skn-1_motif	Oryza sativa	GTCAT	cis-acting regulatory element required for endosperm expression		
	CAT-box	Arabidopsis thaliana	GCCACT	cis-acting regulatory element related to meristem expression		
Other	GCN4_motif	Oryza sativa	CAAGCCA	cis-regulatory element involved in endosperm expression		
Other	HD-Zip	Arabidopsis thaliana	CAAT(A/T)ATTG	element involved in the differentiation of the palisade mesophyll cells		
			CAAT(G/C)ATTG	element involved in the control of leaf morphology development		

Table S3 The putative cis-acting elements and functions predicted by PlantCARE in the $GmRCA\beta$

promoter

Elements type	Names of the elements	Organism	Motif sequence	Function
Light	GA-motif:	Helianthus annuus	AAAGATGA	part of a light responsive element
	I-box	Nicotiana plumbaginifolia	CTCTTATGCT	part of a light responsive element
		Solanum tuberosum	TATTATCTAGA	
	ATCT-motif	Arabidopsis thaliana	AATCTAATCT	part of a conserved DNA module involved in light responsiveness
	Box II	Petroselinum crispum	CCACGTGGC	part of a light responsive element
	Sp1	Zea mays	CC(G/A)CCC	light responsive element
	TCT-motif	Arabidopsis thaliana	TCTTAC	part of a light responsive element
	chs-CMA1a	Daucus carota	TTACTTAA	part of a light responsive element
	GTGGC-motif	Arabidopsis thaliana	CATCGTGTGGC	part of a light responsive element
	Box 4	Petroselinum crispum	ATTAAT	part of a conserved DNA module involved in light responsiveness
	G-Box	Pisum sativum	CACGTG	cis-acting regulatory element involved in light responsiveness
		Daucus carota	TACGTG	
		Arabidopsis thaliana	CACGTG	
Circadian	circadian	Lycopersicon esculentum	CAAAGATATC	cis-acting regulatory element involved in circadian control
Phytohormone	ABRE	Arabidopsis thaliana	TACGTG	cis-acting element involved in the abscisic acid responsiveness
			CACGTG	
			ACGTGGC	
	GARE-motif	Brassica oleracea	AAACAGA	gibberellin-responsive element
Biotic stress	HSE	Brassica oleracea	AAAAAATTTC	cis-acting element involved in heat stress responsiveness
	Box-W1	Petroselinum crispum	TTGACC	fungal elicitor responsive element
Basal element	CAAT-box	Glycine max	CAATT	common cis-acting element in promoter and enhancer regions
		Hordeum vulgare	CAAT	
		Arabidopsis thaliana	CCAAT	
	TATA-box	Arabidopsis thaliana	TATA	core promoter element around -30 of transcription start
		Brassica napus	ATATAT	
		Lycopersicon esculentum	TTTTA	
	5UTR Py-rich	Lycopersicon esculentum	TTTCTTCTCT	cis-acting element conferring high transcription levels
	stretch			
Other	ATGCAAAT motif	Oryza sativa	ATACAAAT	cis-acting regulatory element associated to the TGAGTCA motif
	CAT-box	Arabidopsis thaliana	GCCACT	cis-acting regulatory element related to meristem expression
	O2-site	Zea mays	GATGACATGA	cis-acting regulatory element involved in zein metabolism regulation
	Skn-1_motif	Oryza sativa	GTCAT	cis-acting regulatory element required for endosperm expression

Table S4. The putative cis-acting elements and functions predicted by PlantCARE in the GmRCAa

promoter