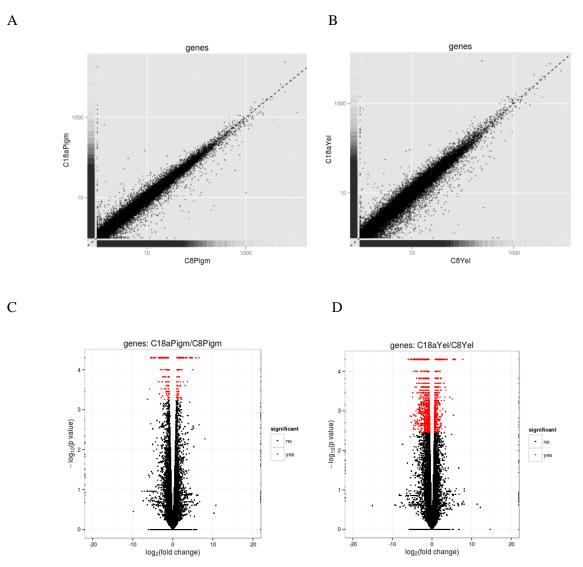


Supplemental Figure 1. The Pedigree of Clark Isolines that Show the Saddle Pattern Phenotype.

The saddle variety having the recessive k1 allele, Clark 18a, is a mutant of Clark which has the dominant K1 allele with i. Another variety showing the saddle pattern phenotype, Clark 8, is an isoline resulting from six generations of backcrossing the i^k allele into the pigmented Clark UC9 isoline which carries the recessive i allele. The seed insets represent the phenotypes. The C or UC numbers are internal lab numbers. The PI numbers are the official USDA germplasm numbers for these lines created by Richard L. Bernard, USDA, in the 1960-70's and are searchable at the national GRIN (Germplasm Resources Information Network) database. The Clark UC9 i mutation (PI547474) is listed as received in 1965 from Portageville, Missouri and is a deletion of *CHS* genes in the I locus region (Todd and Vodkin, 1996). The Clark 18a k1 mutation (PI547439) used in this study is listed as found in Ames, Iowa in 1956.



Supplemental Figure 2. Cufflinks Plots Comparing RNA-Seq Data From Clark (i K1) and Clark 18a (i k1).

These plots assess quality of RNA-Seq using the Cufflinks package (Trapnell et al., 2012).

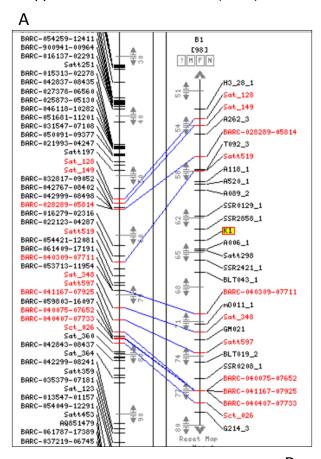
C8: Clark 8; C18a: Clark 18a; Pigm: pigmented region; Yel: non-pigmented region

(A) and (B). Scatter plots highlight general similarities in comparison of gene expression between Clark 8 and Clark 18a (k1) pigmented region (A) or Clark 8 and Clark 18a non-pigmented region (B). Each dot represents an individual gene.

(C) and (D). Volcano plots show the observed differences in gene expression relating to the significance under the statistical model of the Cufflinks package. Genes with an adjusted p value <0.05 are indicated by red dots. There are fewer differentially expressed genes in the comparison of the black saddle regions from Clark 8 (*K1*) to Clark 18a (*k1*) (C) than in the comparison of the yellow sectors (D).

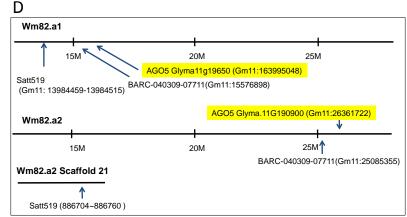
В

С



Object Name	LG	Start cM	Stop cM	Object Type
Sat_128	B1	53.41	53.41	SSR
Sat_149	B1	54	54	SSR
BARC-028289-05814	B1	56.7	56.7	SNP
Satt519	B1	57.91	57.91	SSR
SSR0129_1	B1	61.9	61.9	SSR
SSR2858_1	B1	63	63	SSR
K1	B1	64	64	Gene
Satt298	B1	64.91	64.91	SSR
SSR2421_1	B1	65.7	65.7	SSR
BARC-040309-07711	B1	70.34	70.34	SNP
Sat_348	B1	71.97	71.97	SSR
Satt597	B1	73.76	73.76	SSR
SSR0208_1	B1	76	76	SSR
BARC-040075-07652	B1	77.01	77.01	SNP
BARC-040407-07733	B1	77.01	77.01	SNP
BARC-041167-07925	B1	77.01	77.01	SNP
Sct_026	B1	78.12	78.12	SSR
Sat_360	B1	80.19	80.19	SSR

 $TGACTCTAAAGCAGAACAAGCAAAC [{\bf A} / {\bf T}] TCTACCTTAAATT \\ CCCCAATTCCCAACTTTAAGGCCAGCATTGCTCAATAGAGATC \\ TCCTTCCTTGTGAGTTTCAATTACACAGTCTGCAGTTGGATAG \\ GAGACACAAGTAAGCACAAAGCCTTCCTTCAGTTGGTTGTCAT \\ CGAGGAATGAGCCATCAGACTGGTCCACAGAACCAGCAACAAT \\ CTTCCCAGCACAGGTAGAGCATGCTCCGGCTCGGCAAGAGTAA \\ \\ \end{tabular}$



Supplemental Figure 3. Mapping Information Resources at Soybase for the *K1* Locus.

(A) Screenshot of markers surrounding the position of the *K1* locus (yellow) on Chromosome 11 (historical linkage group B1) as shown at Soybase (Grant et al., 2010) http://soybase.org/ for the soybean composite marker map.

(B) Download of all SSR and SNP

markers surrounding the *K1* gene locus (yellow) between 50 and 80 centimorgans showing two putative flanking markers. Not all markers have available sequences.

(C) Sequence surrounding BARC-040309-07711 used to find its position in the reference genome.

(D) Our diagram of the position of the markers on the two versions of the soybean reference genome (Wm82.a1 and Wm82.a2) showing the relative position of the Glyma model representing the AGO5 gene (yellow) which we identify in this report to be the K1 locus. In the a2 assembly of the genome, only one of the markers (BARC-040309-07711) was located on chromosome 11 while the other marker (Satt519) was found on scaffold 21, possiby due to a discrepancy in the genome assembly.

	1	10	20	17.00162 30	40	50	60	70	80	90	100	110	120	130
H43_protein UC80_protein UC82_protein UC145_protein	MSRRGG MSRRGG	SKQQPDSRHI SKQQPDSRHI	Pappsaapspa Pappsaapspa	GRGRGRGRGRGI	EFSSVPAPSP\ EFSSVPAPSP\	/NAPSAPPPA: /NAPSAPPPA:	STIGYPSGAA STIGYPSGAA	Pyraaspspaa Pyraaspspaa Pyraaspspaa Pyraaspspaa Pyraaspspaa	IESLTSAVEKQ IESLTSAVEKQ	ltlqpsaps Ltlqpsaps	stkavrfser Stkavrfser	PGFGLYGKKI PGFGLYGKKI	kyranhfqyk Kyranhfqyk	vaeqd Vaeqd
UC144_protein UC146_protein UC18a_protein UC150_protein UC151_protein	MSRRGG MSRRGG	SKQQPDSRH SKQQPDSRH	Pappsaapspa Pappsaapspa	GRGRGRGRGRG	EFSSVPAPSP\ EFSSVPAPSP\	/NAPSAPPPA: /NAPSAPPPA:	STIGYPSGAA STIGYPSGAA	Pyraaspspaa Pyraaspspaa Pyraaspspaa Pyraaspspaa	iesltsaveko Iesltsaveko	LTLQPSAPS LTLQPSAPS	stkavrfser Stkavrfser	PGFGLYGKKI PGFGLYGKKI	kvranhfqvk Kvranhfqvk	VAEQD
Consensus								pvraaspspaa 200						
H43_protein		SINPEINSK	VCRDVHTLLV	QAHREKFLG	RIPAYDGRKS	SLFTAGPLPF	ESKDFVIVLK	DEDEPGSSSSA	PARKKREREF	RYTIRFASR	TDLHHLGQFL	RRRQLDCPYE	TIQALDYYLR	ATPSE
UC80_protein UC82_protein UC145_protein UC144_protein UC146_protein	LFHYDV LFHYDV LFHYDV	SINPEINSKI SINPEINSKI SINPEINSKI	VCRDVHTLLV VCRDVHTLLV HTLLV VCRDVHTLLV	QAHREKFLG QAHREKFLG QAHREKFLG QAHREKFLG	IRIPAYDGRKS IRIPAYDGRKS IRIPAYDGRKS IRIPAYDGRKS	SLFTAGPLPFI SLFTAGPLPFI SLFTAGPLPFI SLFTAGPLPFI	ESKDFYIYLK ESKDFYIYLK ESKDFYIYLK ESKDFYIYLK	DEDEPGSSSSA DEDEPGSSSSA DEDEPGSSSSA DEDEPGSSSSA DEDEPGSSSSA DEDEPGSSSSA	iparkkreref iparkkreref iparkkreref iparkkreref	RVTIRFASR RVTIRFASR RVTIRFASR RVTIRFASR	TDLHHLGQFL TDLHHLGQFL TDLHHLGQFL TDLHHLGQFL	RRRQLDCPYE Rrrqldcpye Rrrqldcpye Rrrqldcpye	TIQALDYYLR TIQALDYYLR TIQALDYYLR TIQALDYYLR	ATPSE ATPSE ATPSE ATPSE ATPSE
UC18a_protein UC150_protein UC151_protein Consensus	LFHYDV LFHYDV	SINPEINSKI SINPEINSKI	XVCRDVHTLLV XVCRDVHTLLV	QAHREKFLG QAHREKFLG	IRIPAYDGRKS IRIPAYDGRKS	SLFTAGPLPFI SLFTAGPLPFI	ESKDFVIVLK Eskdfvivlk	DEDEPGSSSSF DEDEPGSSSSF DEDEPGSSSSF DEDEPGSSSSF	iparkkreref iparkkreref	RVTIRFASR RVTIRFASR	TDLHHLGQFL TDLHHLGQFL	RRRQLDCPYE RRRQLDCPYE	TIQALDYYLR TIQALDYYLR	ATPSE
	261 	270	280	290	300	310	320	330	340	350	360	370	380	390
W43_protein UC80_protein UC82_protein UC145_protein UC144_protein UC146_protein UC18a_protein UC150_protein UC151_protein	RFDVVG RFDVVG RFDVVG RFDVVG RFDVVG RFDVVG RFDVVG	RSFFSPFLG RSFFSPFLG RSFFSPFLG RSFFSPFLG RSFFSPFLG RSFFSPFLG RSFFSPFLG	<pre><pgtlgsgtey <pgtlgsgtey="" <pgtlgsgtey<="" pre=""></pgtlgsgtey></pre>	HRGYYQSLRI HRGYYQSLRI HRGYYQSLRI HRGYYQSLRI HRGYYQSLRI HRGYYQSLRI HRGYYQSLRI	PTQMGLSLNIC PTQMGLSLNIC PTQMGLSLNIC PTQMGLSLNIC PTQMGLSLNIC PTQMGLSLNIC PTQMGLSLNIC) VSARAFYEA) VSARAFYEA) VSARAFYEA) VSARAFYEA) VSARAFYEA) VSARAFYEA) VSARAFYEA	LPVIDFIQIH LPVIDFIQIH LPVIDFIQIH LPVIDFIQIH LPVIDFIQIH LPVIDFIQIH LPVIDFIQIH	FRLNPSKPLPD Frlnpskplpd Frlnpskplpd Frlnpskplpd Frlnpskplpd Frlnpskplpd Frlnpskplpd Frlnpskplpd)QDRIKLKRAL)QDRIKLKRAL)QDRIKLKRAL)QDRIKLKRAL)QDRIKLKRAL)QDRIKLKRAL)QDRIKLKRAL	RGIKVEYNH RGIKVEYNH RGIKVEYNH RGIKVEYNH RGIKVEYNH RGIKVEYNH RGIKKLITE	GKNLRRYKIT GKNLRRYKIT GKNLRRYKIT GKNLRRYKIT GKNLRRYKIT GKNLRRYKIT -RILDYTRSL	GYTKEPLREL GYTKEPLREL GYTKEPLREL GYTKEPLREL GYTKEPLREL GYTKEPLREL ESQKNHSES	MFTLDDKR MFTLDDKR MFTLDDKR MFTLDDKR MFTLDDKR	TKSSY TKSSY TKSSY TKSSY TKSSY
Consensus	RFDYYG	RSFFSPFLG	(PGTLGSGTEY	WRGYYQSLRF	PTQMGLSLNIC)VSARAFYEA:	IPVIDFIQIH	FRLNPSKPLPD	DORIKLKRAL	.RGIKve!nh	gknLrrykit;	gvtK#plrel		
W43_protein UC80_protein	ŶQYFHE		ILPALQAGSDS					460 + DRENSIKQVVK DRENSIKQVVK						
UC82_protein UC145_protein UC144_protein UC146_protein UC18a_protein UC150_protein UC151_protein Consensus	VQYFHE VQYFHE VQYFHE	KYNIVLKHTI KYNIVLKHTI KYNIVLKHTI	ilpalqagsds ilpalqagsds ilpalqagsds	KPIFLPYEL(KPIFLPYEL(KPIFLPYEL(QIYAGQRYTK QIYAGQRYTK QIYAGQRYTK	(RLNEEQYTN) (RLNEEQYTN) (RLNEEQYTN)	llratcorph Llratcorph Llratcorph	DRENSIKQYYK DRENSIKQYYK DRENSIKQYYK DRENSIKQYYK	(QSNFSTDKF) (QSNFSTDKF) (QSNFSTDKF)	CHFGIQYKE Chfgiqyke Chfgiqyke	epalldaryl Epalldaryl Epalldaryl	PPPHLKYHGT PPPHLKYHGT PPPHLKYHGT	GRESCYOPRT Grescyoprt Grescyoprt	GQHNM G GQHNM
	521	530	540	550	560	570	580	590	600	610	620	630	640	650
W43_protein UC80_protein UC82_protein UC145_protein	IDKKMY IDKKMY					NLDPLLPIT	SVHSSOVESO	I GNVHKOATAK	I ONECDI EL I	TTTI PNI KG	SYGKTKRTCE	TEL GTYSOCC		QYLEN
UC144_protein UC146_protein UC18a_protein UC150_protein UC151_protein Consensus	IDKKMY	NGGAYQHAT(NGGAYQHAT(NGGAYQHAT(CLNFSGKTNRG Clnfsgktnrg Clnfsgktnrg	ilaasfcqelf ilaasfcqelf ilaasfcqelf	ikmennkgmrf Ikmennkgmrf Ikmennkgmrf	NLDPLLPIT: NLDPLLPIT: NLDPLLPIT:	SYHSSQYESA Syhssqyesa Syhssqyesa Syhssqyesa	LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK	KLANEGRLELL Klanegrlell Klanegrlell Klanegrlell	IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG	SYGKIKRICE SYGKIKRICE SYGKIKRICE SYGKIKRICE	TELGIYSQCC Telgiysqcc Telgiysqcc	lprhyyonkp lprhyyonkp lprhyyonkp	QYLEN Qylen Qylen
UC146_protein UC18a_protein UC150_protein UC151_protein	IDKKMY	NGGAYQHAT(NGGAYQHAT(NGGAYQHAT(CLNFSGKTNRG Clnfsgktnrg Clnfsgktnrg	ilaasfcqelf ilaasfcqelf ilaasfcqelf	ikmennkgmrf Ikmennkgmrf Ikmennkgmrf	NLDPLLPIT: NLDPLLPIT: NLDPLLPIT:	SYHSSQYESA Syhssqyesa Syhssqyesa Syhssqyesa	LGNYHKQAIAK Lgnyhkqaiak Lgnyhkqaiak Lgnyhkqaiak	KLANEGRLELL Klanegrlell Klanegrlell Klanegrlell	IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG	SYGKIKRICE SYGKIKRICE SYGKIKRICE SYGKIKRICE	TELGIYSQCC Telgiysqcc Telgiysqcc	lprhyyonkp lprhyyonkp lprhyyonkp	QYLEN Qylen Qylen
UC146_protein UC18a_protein UC150_protein UC151_protein	IDKKMY IDKKMY 651 I VALKIN VALKIN VALKIN VALKIN	NGGAYQHHTO NGGAYQHHTO NGGAYQHHTO NGGAYQHHTO 660 660 VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI	LNFSGKTNRG LNFSGKTNRG LNFSGKTNRG LNFSGKTNRG 670 + NDAFTRRIPH NDAFTRRIPH NDAFTRRIPH	LAASFCQELA LAASFCQELA LAASFCQELA LAASFCQELA AASFCQELA 680 680 VSDLPTIILO VSDLPTIILO VSDLPTIILO	AKHCNNKGMRF AKHCNNKGMRF AKHCNNKGMRF AKHCNNKGMRF AKHCNNKGMRF ACVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE	NLDPLLPIT NLDPLLPIT NLDPLLPIT NLDPLLPIT NLDPLLPIT SNLDPLT SNLDPL	SVHSSQVESA SVHSSQVESA SVHSSQVESA SVHSSQVESA SVHSSQVESA 710 710 VASMDWPYVT VASMDWPYVT VASMDWPYVT VASMDWPYVT	LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK	KLANEGRLELL KLANEGRLELL KLANEGRLELL KLANEGRLELL KLANEGRLELL KLANEGRLELL REEIIQDLYN IREEIIQDLYN IREEIIQDLYN	ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDYRGK	SYGKIKRICE SYGKIKRICE SYGKIKRICE SYGKIKRICE SYGK 750 750 THSGIIRELL THSGIIRELL THSGIIRELL THSGIIRELL	TELGIVSQCC TELGIVSQCC TELGIVSQCC TELGIVSQCC 760 760 RAFRLSTKTK RAFRLSTKTK RAFRLSTKTK RAFRLSTKTK	LPRHVYQMKP LPRHVYQMKP LPRHVYQMKP LPRHVYQMKP TPRHVYQMKP 770 770 PERIIFYRDG PERIIFYRDG PERIIFYRDG PERIIFYRDG	QYLEN QYLEN QYLEN QYLEN QYLEN 780 1 XSEGQ XSEGQ XSEGQ XSEGQ
UC146_protein UC18a_protein UC150_protein UC151_protein Consensus W43_protein UC80_protein UC82_protein UC145_protein UC146_protein UC18a_protein UC18a_protein UC150_protein UC151_protein	IDKKMY IDKKMY 651 I VALKIN VALKIN VALKIN VALKIN	NGGAYQHHTO NGGAYQHHTO NGGAYQHHTO NGGAYQHHTO 660 660 VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI	LNFSGKTNRG LNFSGKTNRG LNFSGKTNRG LNFSGKTNRG 670 + NDAFTRRIPH NDAFTRRIPH NDAFTRRIPH	LAASFCQELA LAASFCQELA LAASFCQELA LAASFCQELA AASFCQELA 680 680 VSDLPTIILO VSDLPTIILO VSDLPTIILO	AKHCNNKGMRF AKHCNNKGMRF AKHCNNKGMRF AKHCNNKGMRF AKHCNNKGMRF ACVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE	NLDPLLPIT NLDPLLPIT NLDPLLPIT NLDPLLPIT NLDPLLPIT SNLDPLT SNLDPL	SVHSSQVESA SVHSSQVESA SVHSSQVESA SVHSSQVESA SVHSSQVESA 710 710 VASMDWPYVT VASMDWPYVT VASMDWPYVT VASMDWPYVT	LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK 720 720 KYRGYVSAQTH KYRGYVSAQTH KYRGYVSAQTH KYRGYVSAQTH	KLANEGRLELL KLANEGRLELL KLANEGRLELL KLANEGRLELL KLANEGRLELL KLANEGRLELL REEIIQDLYN IREEIIQDLYN IREEIIQDLYN	ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDLKG ITTLPDYRGK	SYGKIKRICE SYGKIKRICE SYGKIKRICE SYGKIKRICE SYGK 750 750 THSGIIRELL THSGIIRELL THSGIIRELL THSGIIRELL	TELGIVSQCC TELGIVSQCC TELGIVSQCC TELGIVSQCC 760 760 RAFRLSTKTK RAFRLSTKTK RAFRLSTKTK RAFRLSTKTK	LPRHVYQMKP LPRHVYQMKP LPRHVYQMKP LPRHVYQMKP TPRHVYQMKP 770 770 PERIIFYRDG PERIIFYRDG PERIIFYRDG PERIIFYRDG	QYLEN QYLEN QYLEN QYLEN QYLEN 780 1 XSEGQ XSEGQ XSEGQ XSEGQ
UC146_protein UC18a_protein UC150_protein UC151_protein Consensus W43_protein UC80_protein UC82_protein UC145_protein UC146_protein UC18a_protein UC18a_protein UC150_protein UC151_protein	IDKKMY IDKKMY 651 I VALKIN	NGGAYQHHTO NGGAYQHHTO NGGAYQHHTO NGGAYQHHTO NGGAYQHHTO S660 VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI VKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI YKYGGSNTYI	LNFSGKTNRG LNFSGKTNRG LNFSGKTNRG LNFSGKTNRG ADAFTRRIPH NDAFTRRIPH NDAFTRRIPH NDAFTRRIPH NDAFTRRIPH NDAFTRRIPH ADAFTRRIPH ADAFTRRIPH ADAFTRRIPH ADAFTRRIPH ADAFTRRIPH ADAFTRRIPH ADAFTRRIPH ADAFTRRIPH ADAFTRRIPH ADAFTRRIPH	LAASFCQELA LAASFCQELA LAASFCQELA LAASFCQELA 680 SOUTING VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO VSDLPTIILO	KMCNNKGMRF KMCNNKGMRF KMCNNKGMRF KMCNNKGMRF KMCNNKGMRF KMCNNKGMRF ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE ADVTHPQPGE	NLDPLLPIT NLDPLLPIT NLDPLLPIT NLDPLLPIT NLDPLLPIT DYSPSIAAV DYSPSIAAV DYSPSIAAV DYSPSIAAV DYSPSIAAV DYSPSIAAV SPSIAAV SPSIAAV SPSIAAV SPSIAAV SPSIAAV SPSIAAV SPSIAAV SPSIAAV SPSIAAV	SVHSSQVESA SVHSSQVESA SVHSSQVESA SVHSSQVESA SVHSSQVESA SVHSSQVESA 710 710 710 710 710 710 710 710 710 710	LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK LGNYHKQAIAK KYRGYYSAQTH KYRGYYSAQTH KYRGYYSAQTH KYRGYYSAQTH	KLANEGRLELL KLANEG	IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG IIILPDLKG SRPTHYRGK SRPTHYRGK SRPTHYRGK SRPTHYRUL SRPTHYRUL SRPTHYRUL SRPTHYRUL	SYGKIKRICE SYGKIKRICE SYGKIKRICE SYGKIKRICE SYGK 750 750 750 750 750 750 750 750 750 750	TELGIVSQCC TELGIVSQCC TELGIVSQCC TELGIVSQCC 760 RAFRLSTKTK	LPRHVYQMKP LPRHVYQMKP LPRHVYQMKP LPRHVYQMKP LPRHVYQMKP T770 770 PERIIFYRDG PERIIFYRDG PERIIFYRDG PERIIFYRDG PERIIFYRDG PERIIFYRDG PERIIFYRDG PERIIFYRDG TYRCTRSVS TYARCTRSVS TYARCTRSVS TYARCTRSVS	QYLEN QYLEN QYLEN QYLEN QYLEN 780 780 YSEGQ YSEGQ YSEGQ YSEGQ YSEGQ YSEGQ YSEGQ YSEGQ YSEGQ TSEGQ TSEGQ TSEGQ TSEGQ TSEGQ TSEGQ TSEGQ TSEGQ TSEGQ TSEGQ TSEGQ TSEGQ TSEGQ

Supplemental Figure 4. Alignment of Nine AGO5 Proteins Derived from Amplicon Sequences Comparing *K1* and *k1* Lines. (continued from previous page)

The cultivar or mutant genotypes of the following nine predicted proteins derived from amplicon sequencing were aligned with Multalin http://multalin.toulouse.inra.fr/multalin/

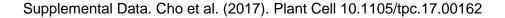
W43	PI548631	Williams	Reference	K1
UC80	PI548362	Lincoln	Parent line	K1
UC82	PI548527	Calland	Parent line	K1
UC145	PI634895	L88-5422	saddle pattern mutant from Lincoln 1945	k1
UC144	PI634873	L88-5344	saddle pattern mutant from Calland 1970	k1
UC146	PI634896	L88-5424	saddle pattern mutant from Lincoln 1954	k1
UC18a	PI547439	L67-3479	saddle pattern mutant from Clark 1956	k1
UC150	PI506948	Kurakake	saddle pattern variety from Japan, 1987	k1
UC151	PI506949	Kurakake Daizu	saddle pattern variety from Japan, 1986	k1

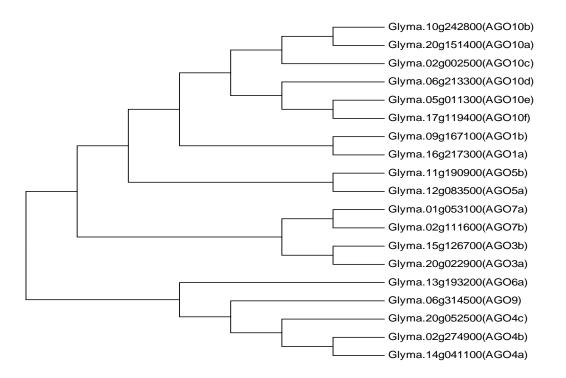
	1	10	20	30	40	50	60	70	80	90	100	110	120	130
Glyna.11G190900.1 Glyna.12G083500.1_AG Consensus	ATGTCT	CCTCGCGG1	rggctc <mark>caag</mark>	CAGCAACCA	GATTCACGCCA GATTCACGCCG GATTCACGCCa	TCAAGCGCC	ACCGTC <mark>C</mark> GCG(i <mark>CG</mark> GCGGCGG	TGCCTTCTCCT	CCCGGAAGA	GGACGCGGCC	GTGG <mark>A</mark> CGCGG	AGAATCC	TC <mark>r</mark> gttc TCCGTTC
Gluma.116190900.1	131 		150	160	170 CTCCTCCGGCT	180		200	210	220	230	240	250	260 1
lyna,12G083500,1_AG Consensus	CCGCGC	CTTCTCTG	GTGAATGCAC	CATCTGCTC	CTCCTCCGGCT	TCCACCATC	GGCGTTCC <mark>G</mark> T(CGGCGCCGC	GCCTTTTCGT	ictgcgtt <mark>g</mark> c	CACCTCCGGC	GGTGGAATCT	CTCACTTCCG	C <mark>g</mark> gtcga
61yma.116190900.1	261 	270		290	300 GTCGACGAAGG	310	320	330		350	360	370		390 0010001
lyma,126083500,1_AG Consensus	AAAGCA	ACTTACA <mark>a</mark> t	FGCAACCTTC	GGCGCC <mark>G</mark> TC	GTCGTCGAAGG GTCG <mark>a</mark> CGAAGG	CGGTGAGGT	tca <mark>ag</mark> gaacg(iccagggttc	GGTCTTG <mark>C</mark> TGG	iggagaaaat	CAAAGTTCGA	GCGAATCATT	TTCAAGTCCA	AGTCGC <mark>C</mark>
Glyma.11G190900.1	391 	400			430 HATCC <mark>g</mark> gagat	440	450	460		480	490	500	510 	520 1 CCGGCCT
lyna,126083500,1_AG Consensus	GAACAA	GATCTATT	CATTATGAC	GTCTCTATC	AATCCTGAGAT AATCCgGAGAT	TACTTCGAA	GAAGGTTT <mark>CT</mark> A	IGAGATGTTA	TGACTTTGCTT	GT <mark>G</mark> CAAGCG	CATCGTGAAA	AGATTCTCGG	AAATCGCATA	CCAGCCT
Gluma.116190900.1	521 	530 +	540 GTCTTTTC9	550	560 CCTTGCCCTTC	570	580 +	590	600 +	610 +	620	630	640	650 ААСБТБА
lyna,126083500,1_AG Consensus	ACGATG	GGGGAAAGA	AGTCTTTTCA	CCGCTGGAT	CCTTGCCCTTC	GAATCTAAG	GATTTTGTGA	CGTGCT <mark>a</mark> aa	AGACGACGATO	AACCAGGCT	CGTCGTCTTC	TTCCTCTCCT	<mark>a</mark> ctaggaaaa	AACGTGA
Cluma 110100000 1	651 	660	670	680	690 + CAGGACTGACC	700	710	720	730	740	750	760	770	780 1
Glyna,11G190900,1 ilyna,12G083500,1_AG Consensus	ACGCGA	GT <mark>a</mark> tagagi	FCACCATCAG	G <mark>cttgcttc</mark>	CAGAACTGACA CAGAACTGACA CAGaACTGACa	TTCACCACC	TC <mark>A</mark> GTCAGTT1	CTCAGGCGC	CGTCAGTTGGF	ITTGTCCTTA	TGAGACTATC	CAGGCTCTTG	ATGTTGTTCT	GCGTGCT
C1 440400000 4	781 	790	800	810	820	830	840	850	860	870	880	890	900	910 I
Glyna,11G190900,1 ilyna,12G083500,1_AG Consensus	ACACCG	TCTGAAAGO	GTTCGTTGTT	GTGGGAAGA	TCGTTCTTCTC TCGTTCTTCTC	ACCTTCTTT	GGGGAAACCT(iGATCGCTTG	GTAGCGGAACO	igagta <mark>c</mark> tgg	AGGGGCTATT	ACCAGAGCCT	tcg <mark>c</mark> ccaact	CAGATGG
Glyna.116190900.1	911 	920	930	940	950	960	970	980	990	1000				1040
lyna,126083500,1_AG Consensus	GCCTGT	CTCTTAACA	ATT <mark>A</mark> ATGTGT	CGGCGAGGG	CTTTTTATGAG CgTTTTATGAa	CCTATTCCT	GTGATTGATT I	CATTGAAAG	TCATTTTAGGE	CCAATCCTT	CCAGGCCTTT	GCCTGATCAG	GATCGAATCA	AGCTTAA
	1041 	+	1060	1070	1080	1090	1100	1110	1120	1130	1140	1150	1160	1170
Glyna,116190900,1 lyna,126083500,1_AG Consensus	GAGAGT	A <mark>c</mark> tgagag(GAGTGAAGGTI	AGAAGT <mark>G</mark> AC	TCA <mark>C</mark> GGAAAGA TCATGGAAAGA TCA <mark>c</mark> GGAAAGA	ATCTTAGAC	GTTACAAGAT(ACTGGAGTC:	acaaaagaac <mark>f</mark>	ACTCAGAAA	GTTAATGTTT	ACTCTTGATG	ACAATAGAAC	AAAAAGC
	1171 	+	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300
61yna,116190900,1 1yna,126083500,1_AG Consensus	TCAGTT	GTTCAATAT	FTTTCATGAG	RAATACAAT	ATTGTGTTGAA Attgtgttgaa Attgtgttgaa	GCATACGCT	TCTTCCTGCT(TTCAAGCTG	GTAGTGACATI	AAACCAATT	TTT <mark>C</mark> TGCCT <mark>r</mark>	T <mark>g</mark> gagctttg	tcaaattgt <mark>g</mark>	GCTGGAC
	1301 	+	1320	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430 l
Glyna,116190900,1 lyna,126083500,1_AG Consensus	AAAGAT	ATACAAAGA	IGATTGAATG	AGGAGCAAG	TAACTAATCTT TAACTAATCTT TAACTAATCTT	TTAAGGGCA	TCTTGTCAGCO	ITCCTCGTGA	TAGAGAAAACT	CTATCAGAC	AGGT <mark>G</mark> GTGAG	GCAAAGTAAT	TTCAGCACAG	ACAAATT
		+		1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560 I
Glyna,116190900,1 lyna,126083500,1_AG Consensus	TGTG <mark>a</mark> g	TCACTTTG	GAATTCAAGT	Gagggagga	ACCAGCATTGC TCCAGCATTGC aCCAGCATTGC	TTGATGCTC	Gagttcttcc 1	GCACCAATG	CTAAAATATCF	ITG <mark>a</mark> cacagg	TAGAGAATCA	AGTGTTGAAC	CCA <mark>R</mark> AATGGG	CCAATGO
	1561 	1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690
Glyna.116190900.1 lyna.126083500.1_AG Consensus	AATATG	ATTGATAAC	GAAAATGTTTI	AATG <mark>c</mark> tggt	G <mark>C</mark> TGTG <mark>C</mark> AGCA GTTGTGGAACA GcTGTGcAaCA	TTGGACTTG	CCTCAACTTT	CTGGAAAAA	TAAACAGAGAG	TTTCCAAGT	GCATTTTGTC	AT <mark>R</mark> AGTTGGC	CAGAATGTGT	AGTAACA
		+	1710	1720	1730	1740	1750	1760	1770	1780	1790		1810	
61yna,116190900,1 1yna,126083500,1_AG Consensus	AGGGCA	TGCGTTTTF	AATT <mark>CAAAG</mark> C	CTTT <mark>a</mark> ctgC	CCATAACATCT CCATAACATCT CCATAACATCT	GCTCARAGT	agtcaa <mark>a</mark> tagi	IGAGTGCTCT	TGTAAATTTGC	:ATAAGCAGT	CTATTACAAG	ACTAGCAAAC	CAAGGAAGAC	TCCAATT
	1821 	1830 +	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950
Glyna,116190900,1 lyna,126083500,1_AG Consensus	GTTGAT	CATAATTTI	FGCC <mark>A</mark> GATTT	CGAGGGGGTC	CTATG <mark>GG</mark> AAAA CTATG <mark>AA</mark> AAAA CTATG <mark>aa</mark> AAAA	TAAAGCGTA	TTTGTGAAAC1	GAG <mark>C</mark> TAGG <mark>A</mark>	ATAGTGTCTCF	IGTGTTGTC <mark>A</mark>	GCCGAGGCAT	GTTTGCCAGA	TGAAGCCACA	ATATCTT
		+	1970	1980	1990		2010			2040	2050	2060	2070	2080 l
Glyna,11G190900,1 ilyna,12G083500,1_AG Consensus	GAAAAT	GTGGCCCTO	CAAGATAAAT	GTGAAGGTT	GGTGGCAGTAA GGTGGCAGTAA GGTGGCAGTAA	CACAGTATT	GAATGATGCA	TTGCTAGAA	TAATTCCTCGT	GTGTCTGAC	AGACCTACAT	TAATCTTGGG	TGCGGATGTA	ACACATO
		+			2120							2190		
Glyna,116190900,1 lyna,126083500,1_AG Consensus	CACAGC	CAGGGGAAC	GATTCTAGTC	CTTCTATTG	CTGCAGTAGTT CTGCAGTAGTT CTGCAGTAGTT	GCATCTATG	GATTGGCCTTA	ITGTAACAAG	GTACAGAGGAG	ITTGTTTCTG	CTCAGACTCA	CCGTGAAGAA	ATCATCCAAG	ATCTTT
	2211	2220	2230	2240	2250	2260	2270	2280	2290	2300	2310	2320	2330	2340

Supplemental Data. Cho et al. (2017). Plant Cell 10.1105/tpc.17.00162														
61yna.116190900.1 61yna.126083500.1_AG Consensus	TAATACA	rgtgaaq	GATCCTGTG	AGGGGAAGG	CACATTCGGGF TGCATTCGGGF caCATTCGGGF	ATTATCAGG	GAGTTACTTC	GCGCTTTCCG	TTTGTCTACTE	iatc <mark>ag</mark> aagci	CAGAGAGGATI	FATATTCTACA	GGGATGGAGT	AAGTGAG
	2341 23	350	2360	2370	2380	2390	2400	2410	2420	2430	2440	2450	2460	2470
61yna,116190900,1 Glyna,126083500,1_AG Consensus		FCAGCCF	AGGTTTTGC	GTACGAGAT	GGATGCAATAO GGATGCAATAO GGATGCAATAO	GGCGGGCTT	GTGCTTCACT	ACAAGAAGGC1	TATTTACCCCC	TGTTACTTT	IGTGGTGGTCC		CCACACAAGG	
	2471 24	480	2490	2500	2510	2520	2530	2540	2550	2560	2570	2580	2590	2600
Glyna,11G190900,1 Glyna,12G083500,1_AG Consensus	CTGTAGA	TCATGGE	AGTCATGA	CAGACAAAT	'AAAAGTGGAAA 'AAAAGTGGAAA 'AAAAGTGGAAA	ITATA <mark>A</mark> TGCCI	AGGGACTGTC	gt <mark>a</mark> gacac <mark>a</mark> ca	ACATATGCCAC	CCTCGGGAG	ITTGATTTTT	ACCTCAACAGT	CATGCTGGAA	
	2601 20	510	2620	2630	2640	2650	2660	2670	2680	2690	2700	2710	2720	2730
61yna,116190900,1 Glyna,126083500,1_AG Consensus	AACTAGT	CGACCAF	ICACATTAT	ATGTGCTGT	TCGATGAAAT TCGATGAAAAA TCGATGAAAa	AACTTCACT	GCTGACGGGT	FGCAAATGTT	ТАСТААТААТІ	TGTGTTATA				
	2731 23	740	2750	2760	2770	2780	2790	2800	2810	2820	2830	2840	2850	2860
Glyna.11G190900.1 Glyna.12G083500.1_AG Consensus	CCTGTGT	ATTATGO	CACATTTGG	TGCCTTCAG	GGCTCGCTGTT	ACATTGAAG	TGCAACATC	AGATTCTGGT	TCTGCAAGTGO	AGGCCGGGC	rgctaactgto	GAGGTTAGATT	GCCTTCGGTT	aaggaaa
	2861 20		2880	2889										
Glyna,116190900,1 Glyna,12G083500,1_AG Consensus			GATGTTTTT	TGTTAA										

Supplemental Figure 5. Alignment of Two Closely Related Soybean AGO5 Coding Sequence Transcripts.

The coding sequences of Glyma.11G190900.1 (2673 bp) and Glyma.12G083500.1 (2886 bp) were aligned using Multalin at http://multalin.toulouse.inra.fr/multalin/.





Supplemental Figure 6. Phylogenetic Tree for Soybean Argonaute Genes

(Top) Phylogenetic analysis for 20 soybean AGO genes was performed with MUSCLE and MEGA v6.0 programs by the maximumlikelihood method with 500 bootstrap replicates as Liu et al. (2014) described.

(Right) The gene model number of AGO genes in both soybean reference genomes. The list of Wm82.a1.v1.1 was reported in Liu et al., 2014 and the a, b, c, etc. subclassification follows that nomenclature to distinguish homologs.

Annotation	Wm81.a1.v1.1	Wm82.a2.v1
AGO1a	Glyma16g34300	Glyma.16G217300
AGO1b	Glyma09g29720	Glyma.09G167100
AGO3a	Glyma20g02820	Glyma.20G022900
AGO3b	Glyma15g13260	Glyma.15G126700
AGO4a	Glyma14g04510	Glyma.14G041100
AGO4b	Glyma02g44260	Glyma.02G274900
AGO4c	Glyma20g12070	Glyma.20G052500
AGO5a	Glyma12g08860	Glyma.12G083500
AGO5b	Glyma11g19650	Glyma.11G190900
AGO6a	Glyma13g26240	Glyma.13G193200
AGO6b	Glyma15g37170	no correspondence
AGO7a	Glyma01g06370	Glyma.01G053100
AGO7b	Glyma02g12430	Glyma.02G111600
AGO9	Glyma06g47230	Glyma.06G314500
AGO10a	Glyma20g28970	Glyma.20G151400
AGO10b	Glyma10g38770	Glyma.10G242800
AGO10c	Glyma02g00510	Glyma.02G002500
AGO10d	Glyma06g23920	Glyma.06G213300
AGO10e	Glyma05g08170	Glyma.05G011300
AGO10f	Glyma17g12850	Glyma.17G119400
AGO10g	Glyma04g21450	Glyma.04G150900