

Fig. S1. Native genomic *rhg1-a*  $\alpha$ -SNAP amplicons from PI 89772 are unexpectedly large and contain an inserted DNA element. (A and A') Agarose gel showing PCR amplification of the native genomic *rhg1-a*  $\alpha$ -SNAP amplicon, using a forward primer 0.85 kb upstream of transcript start and a reverse primer within region encoding 3'UTR. Similar results obtained from two independent genomic (gDNA) DNA preparations of PI 89772 and with varying primer annealing temperatures. (B) Agarose gel of PCR products using primers "RAC amplify" F & R on genomic DNA from PI 89772 (*rhg1-a*), Williams 82 (*Rhg1*<sub>WT</sub>, single copy at *Rhg1*), and Fayette (*rhg1-b*), again showing additional 4.7 kb of DNA within *rhg1-a* but not wild-type *Rhg1* or *rhg1-b*. (C) Agarose gel showing no detection of RAC- $\alpha$ -SNAP junctions within the WT *Rhg1*  $\alpha$ -SNAP encoded by the soybean reference genome, Williams 82 (Wm82), relative to a positive control plasmid containing the subcloned PI 89772 (*rhg1-a*) RAC- $\alpha$ -SNAP (+RAC Ctrl); H<sub>2</sub>O - no template negative control. (D) Schematic of *Glyma.18G022500* within *Rhg1* locus (copia is only present in *rhg1-a* alleles), showing sites of homology for PCR primers used in (B) and (C).

## $\alpha$ -SNAP<sub>Rhg1</sub>LC RAC (*rhg1-a* associated copia) Nucleotide Sequence

### Features:

$\alpha$ -SNAP<sub>Rhg1</sub>LC Exon 1

$\alpha$ -SNAP<sub>Rhg1</sub>LC Intron 1

RAC 3' LTR

RAC Polyprotein ORF (anti-sense to  $\alpha$ -SNAP<sub>Rhg1</sub>LC ORF)

RAC 5' LTR

$\alpha$ -SNAP<sub>Rhg1</sub>LC Exon 2

Primer binding site

Putative polypurine tract (3'-GAGGGGGG-5')

```
ATGGCCGATCAGTTATCGAAGGGAGAGGAATTTCGAGAAAAAGGCTGAGAAGAAGC
TCAGCGGTTGGGGCTTGTTGGCTCCAAGTATGAAGATGCCGCCGATCTCTTCGAT
AAAGCCGCCAATTGCTTCAAGCTCGCCAAATCATGTTTTCTCTTTCTCTCTACTT
TTTTTAAATTCCATTCGTGTCTCCTCAAATGTTGATTTAGTGTCAATAATCATAATT
ATTATTCTCTTCTATTGTTGTTATTTTATTGTTATTACTTCAATCGACGAGTGTGTTGA
GTTTTGAGGTGTCCGATTTCCCGATTAATTGAAGTATAGTTTTAATCTGATTTTACTG
GAAAATATTTTTTGCCTGATTTTTGTTTTTGGAAACAATACTAGCATATAAATTAGA
ATTGTGGATGAAGTATAAGAAACTACAGGAGCATAAAGGAAGAAGAAGTGAGCTTG
AATATTTTCAGAGAAGAAGATCAGCTTCAGCTACTTATTTTCGTATTAACAGAGAAGG
TTTATATACATGATGTGTGATTGTTATAACAGAAAAGCTAACTAACTCAACTAACCC
AACTACCCTTAACTGATACTGTTATACTGCTAAGAGCCCCCTCAAGCTGGGAATG
GATATTCATCATTCCCAGCTTGTTACAGAGATGCCGAAAAACAGCAGGTGCAAGAG
CTTTGGTGAATATATCCGCTAGTTGTAAGCTGACGAAACCGGAAGAAGCTTTAGG
AGACCTGAGTTAAGCTTTTGACGAACAATATGACAGTCTATCTCGATATGTTTAGTT
CGTTCGTGAAAAACGGGATTAGTAGCTATTTGGATGGCTGACTGATTATCACAGTA
TAAGTTCGCTGGTTGAACGAATGAGATACGAAAGTCTTGAAGCAAGAAGGTCAGCC
ACTGTAGTTCGCAAGTAGTGGAGGCAAGAGCGCGGTATTCAGCTTCGGAAGAACT
GCGTGACACAGTTGGCTGCTTCTTGGATTGCCAAGAAACCAGAGAGGAACCAAAG
TAACTAAGTAGCCGGTAGTGGATTTCCTTGAATCTTTGCATCCAGCCCAATCCGA
GTCATAAAGGCTCGGAGTTGTGCGGTACCTGCGGCAGTGAAGAAGATACCTGAT
CCCGGAGAACTCTTGAGGTATCGAAGAATCCGAAAGGCGGCTTGAAGATGAGCAT
TGGTGGGGGCGGCCATGTACTGGCTGAGTTGTTGAACAGCATAACGTTATATCGGG
CCTGGTGTGGTAAGGTATATTAATTTACCGATCAATCGCCGATAAGAGGAAGAAG
ACTCAGCTGAGAGAGGACTGCCCGAATCTGCCTGTAACCTTCGTAGAGTAGTCTATT
GGTGTGTAATTGGGCTTGCATCCAGCATTCCGGATGCATTTAGAATGTCTAATGT
ATACTTGC GTTGGCATAAGTGTATCCCTTTCGAGCTTCGGGCGATTTCAAGCCCA
GGAAAACTTTAAATCCCAAGATCCTTGATCTTAAATTCAGAATCTAAGAGGGTGA
CAATTGTTGTATTTCGGTCATGCTATTTCTGTGAGAATGATGTCGTCTACATAAA
CAAGAAGGATGTTGTGATGTTTCCAGTAAACCGCAAAAAGAGAGAGTGATCCGC
```

AGTTGACTGATGAAAGCCATGAGAGGTTAAGAAGCTTGACAATTTTACGAACCATT  
GTCTGCTGGCTTGTTTGAGACCATAAAGAGACTTTTGAAGGCGACATAACAAGCTTT  
GGGTTATCAACGGAAAGTCCCGGAGGTATTTGCATATAAACCTCCTCGTCAAGTTC  
TCCATGGAGGAATGCATTATTAACATCTAGCTGCCGTAATGCCACTGGTTTAATG  
CTGCAATTGCAAGAAGAAGGCGCACCGTGGTCAGCTTTGCTACCGGAGAGAAAGT  
GTCAAGGTAGTCTAACCCCTTCATTTGGGTGTATCCCTTTGCAACCAGCCGCGCTT  
TATGCCTTTTCGATGGATCCATCTGCTCTATACTTTATTTTATAGACCCATCTGCACC  
CAATAGCCGTCTTGTGAGAAGGGAGAGGTGTGAGGCGCCATGTTTGGTTCGACTG  
AAGAGCTCGTAGCTCGGCTTCCATGGCCTTAATCCAGCAATCATGGCGAGAAGCA  
TCGACATATGAGGTTGGCTCTGTGACGGAGGAAATATTCATGACAAAGTTCCTGTG  
GGCAGGAGACAAGCGTGAGTAAGAGAGTACGGAACTAAGTGGATAACGAACAGCC  
ATTGAAGTGCTTGGTGTAGAGGAAGCAAACCTCTCTGTGGTAATCTCGGAGGTACGT  
TGGGGTGTTTTTGGTTCTGGTGGATCGTCTAAGGTGTGAAGGAGGATGATTATGTT  
CAGGTTCAATTTGATGATGGTATAGAGATCATAGGTGGTGTATGAAAGACTCTCTGTTT  
GTGGGTCATCGTTTCTGTTCGGAGAAGGATTCCGGTGAGGGAGCTGGATATTCTAA  
GTGTGTATGCTGAGTTTCAGAAAGATAAGGAAAATGATCCTCATAAAATGTGATATT  
TCGAGAGATGCTAACATCATTAGAGTGCAAATCATAACAAGATATCCCTTTGTATG  
CATTTTAAAACCGATGAATATGCATGGATGAGCCTTAGCATCAAGCTTTTGCCGGTT  
TGCCTTGAGTGTATTTATGTAACATAGACACCCGAAAACACGAAGGTTAGAAATGT  
CACAAGGGTGTATTATGCAGCTTTTCATAGGGTGAAACATTATGCAAAAACGGCGTG  
GGAATACAATTAATCAAGTAAGTGGCATGCGGCAAAGCGTAACACCAGAAGCTTG  
GTGGTAGACTTGCCTGAAACAAAAGTGCACGTGTGACATTGAGAAGGTGCTGGTG  
TTTGCCTTCTACAATTCCGTTTTGTTCTGGAGTTTCAATGCACGTGGTCTGGTGTAT  
GATGCCCTTTGATGCATAGTAATGATGCATGGAGAATTCATGCCATTATCACTTCT  
GATGATCTTAACCTTGCCATCGTATTGTGTTTCAATGAATGTAATGAAGTTCATGAT  
TATATGTCGGGTTTCAGCTTTGGATTTCATAAGATGAACCCATGTAAAGCGTGAGC  
AATCATCAACTATAGTTAAGAAGTATTTGTGCCATGCATGGATGGTTTAGAGCAC  
GGACCCCATATGCCATATGCAGTAAGTCAAAAATGTGAGATGCATGTGAATGGCT  
AAGAGAAAAGGTAATTTCTTTTGTTCGCATGATGACACGTGTTGCAAACAAAATC  
CTTATTATTTTTGAGAAGGGGATAGTAAGCTTTACATACATTGTATTCTTTTCAAGTGGAT  
GGGTGGCCTAACCTAAAATGCCAAAGGTCAATAGGTATTACATTACATCGAGGGTG  
AGTAATAGTGGAGTTTACGGTTTTGGTGGTCAGCTGAGCAGGTATTAATGGTAGA  
GACCGTGTTTTGTTCACCTATAACCAATCCTCATATGGCTGTTCACTTCTCTGTAATA  
CACACGATGTAGAGGAGAATATCAATTCACAATTAACGGAAGACACAAGTTTTGAT  
ATTGAAATGATATTGAACGTAAAGGAAGGAATGTAAAGAACGTCTTGTAATGATG  
TTTGATGAAAGCTTGACAATCCTGAGTGGGTTGCACAAACACACTGGCCATTCGG  
GAGTTTACCGTGATAGGATCGATTTGTTTATAGGAATGAAGGTTGCGTAGGGAGT  
AAGTCGCGTGGTCCGTTGCTCCTGAATCCAATATCCAGGAGACGCATGGTGTCT  
GAGAGAAAGGGACATACCTGGATTTGTCGGAGTATTGATACATGATGAGATTGAAG  
CCATTTGTTTGGATTGAGAGATTGTCGTGTTTCTGCGGATGGTTCCTGGATTAAG  
GCTAGGAGTGCCTTGTACTGCTCGGGGGAGAAACGAACAGAATCATGAGCCTCGT  
GGTGCTGAGCTTGATCTTCGTTAGCTTTGTTTTCAACGGCTACGAGATTGTTTACAG  
TGGTTCTTCCATAGGGCTTGTAAACCGGCGTGTAAACCGGTGGGTACCCGTG

TTTTCGATAACACACAGTCCACAGTGTGTCCCATCTTGCCACAATGAGTGCAAGCTT  
TCCTTCCACTATTCTTGTTTCTAGTATCGTGATTAGAGGGCATTCCATGTTTCTTATA  
ACATGTGCTTTCTAAGTGACCAACACGTCCACAGAAGTCGCAGACAGTTTTAGCAG  
CGTTTATAGAAATTTCTTTGGGTTTGAAATGAATACCTGGTCCAGCGTTTCCCAGTA  
ATTGCCTTTCCTGTTGAGCCACATAGGAGAAAATCTTGGAGATAGCAGGTATAGGA  
TCCATGAGAAGAACGTGGGATCGTATATTTCCGTA CTGTTTATTGACACCGCGTAG  
AAACTGCATGGCCCTATCTTCGAGTTTCCGCTGTGCGATAATGGTGAATGCATTGC  
AGGAACATCTTATATTACATGAGCAAATGGGATCGGGTCTAAAGTTCTCGATTTCGT  
CCCAAATGACGCGTAATCGTGTGAAATACTCAGTTACTGTGAGCGTACCTTGCTTC  
ATCGTCGAAGCTTCTTGTTGAAGGTCGGATATGCGTAAAAGATCTCCCTGAGAGTA  
TCTTGATTTAGATCGCGCAAATTTCTCGGCTTTGTCCATCCAAAGTATGCTCTG  
ACGTATGGAGATGGCCACCGAATGAACTATCCACGAGACGACCATATTGTTACATC  
TACGCCATGCTCCGTGCATTCTGTCCGTTTTTAGAGGTTCCGGGGCGCTGCCATCT  
ATGAACTCCACTTTATTCTTGGCACTCAATGCAGTGACCATAGACCTGCTCCATGA  
GTGGTAATTGGTTGAGTCTAGGACTGGGGAAACAAGAGCGATAGCTGGGTTTTTCG  
CTTGATGGAGGTAGAGATAACTCTCCATGTTACTAGCAGAAGATTCGTTTCATGGT  
GGATAATGGAAGAATGCGCAGCAGAACTCTTCTTTT **AGAAGAGCTCTGATACCA**TA  
AGAAAACACAGGAGCATAAGGAAGAAGAAGTGAGCTTGAATATTTAGAGAAAGAA  
GATCAGCTTCAGCTACTTATTTTCGTATTAACAGAGAAGGTTTATATACATGATGTG  
TGATTGTTATAACAGAAAAGCTAACTAACTCAACTAACCCAACTACCCTTAACTGAT  
ACTGTTATACTGCTAAGA **AAGTACGACAATCAACTCTGTGTTGTTTGTGACTACGCT**  
CACTTTCAATTTGACGACTAATCTCTTTATTTTGTGAAAGTGACGAACTTTGAAATT  
GATGTTGGAATAGTTCTGTTTATTGTTCTTGATTTGATCTATGTGGCATTTTAG **GGG**  
**ACAAGGCTGGAGCGACATACCTGAAGTTGGCAAGTTGTCATTTGAAG**

Fig S2. Complete nucleotide sequence of the PI 89772 (*rhg1-a*, *Rhg1* low-copy) *RAC* (element and flanking exonic  $\alpha$ -*SNAP*<sub>*Rhg1*</sub>*LC* regions. Key sequence features are color coded as indicated in the above box.

### S3

#### PI 89772 *rhg1-a* RAC-encoded polyprotein (1438 residues)

MESYLYLHPSENPAIALVSPVLDSTNYHSWSRSMVTALSAKNKVEFIDGSAPEPLKTDR  
MHGAWRRCNNMVVSWIVHSVAISIRQSILWMDKAEEIWRDLKSRYSQGDLLRISDLQQ  
EASTMKQGTTLVTEYFTRLRVIWDEIENFRPDPICSCNIRCSCNAFTIIAQRKLEDAMQ  
FLRGLNEQYGNIRSHVLLMDPIPAISKIFSYVAQQERQLLGNAGPGIHFEPKEISINAAKT  
VCDFCGRVGHLESTCYKKHGMPSNHDTRNKNSGRKA**CTHCGKMGHTVDVC**YRKHG  
YPPGYTPGYKPYGGRTTVNNLVAVENKANEDQAQHHEAHDVSRFSPEQYKALLALIQ  
EPSAGNTTISQSKQMASISSCINTPTNPGMSLSLRTPCVS**WILDSGATDH**ATYSLRNLH  
SYKQIDPITVKLPNGQCVCATHSGIVKLSSNIILQDVLYIPSFTFNIISIKLVSSVNCCELIFS  
STSCVLQEVNSHMRIGIVEAKHGLYHLIPAQLTTKTVNSTITHPRCNVIPIDLWHFRLGH  
PSTERIQCMKAYYPLLKNNKDFVCNTCHHAKQKKLPFSLSHSHASHIFDLLHMDIWGP  
CSKPSMHGHKYFLTIVDDCSRFTWVHLMKSKAETRHIIMNFITFIETQYDGKVKIIRSDN  
GIEFSMHYYASKGIIHQTTCIETPEQNGIVERKHQHLNVTALLFQASLPPSFWCYAL  
PHATYLINCIPTPFLHNVSPYEKLHKHPCDISNLRVF**GCLCYINTLKANRQKLDARAHPC**  
**IFIGFKMHTKGY**LVDLHSNDVVISRNITFYEDHFPYLSETQHTHLEYPAPSPESFSDRN  
DDPQTESLSSPPMISIPSSNEXEHNHPPSHLRRSTRKNTPTYLRDYHREFASSTPSTS  
MAVRYPLSSVLSYSRLSPAHRNFVMNISSVTEPTSVDASRHDCWIKAMEAELRALQS  
NQTWRLTPLPSHKTAIGCRWVYKIKYRADGSIERHKARLVAKGYTQMEGLDYLDTFSP  
VAKLTTVRLLLAIAALNQWHLR**QLDVNNAFLHGELDEEVYMQIPP**GLSVDNPKLVCRLQ  
KSL**YGLKQASRQW**FVKLSSFLTSHGFHQSTADHSLFLRFTGNITTILL**VYVDDI**IILTGNS  
MTEIQTIVTLLDSEFKIKDLGDLKFFLGLLEIARSSKGIHLCQRKYTLDILNASGMLGCKPN  
STPIDYSTKLQADSGSPLSAESSSSYRRLIGKLIYLTNTRPDITYAVQQLSQYMAAPTNA  
HLQAAFRILRYLKSSPGSGIFFTAAGTAQLRAFSDSDWAGCKDSRKSTTGYL VYFGSS  
LVSWQSKKQPTVSRSSSEAEYRALASTTCELQWLTFLLQDFRISFVQPANLYCDNQSA  
IQIATNPVFHERTKHIEIDCHIVRQKLNSGLLKLLPVSSALQLADIFTKALAPAVFRHLCNK  
LGMMNIHSQLEGG\*

#### Conserved Structural Features of Copia Polyproteins

GAG binding motif: **CTHCGKMGHTVDVC**

Protease motif: **WILDSGATDH**

Integrase (GKGY motif): **GCLCYINTLKANRQKLDARAHPCIFIGFKMHTKGY**

Rev. Transcriptase motif:

**QLDVNNAFLHGELDEEVYMQIPP**GLSVDNPKLVCRLQKSL**YGLKQASRQW**FVKLSSFL  
TSHGFHQSTADHSLFLRFTGNITTILL**VYVDDI**I

Fig. S3. Translation of the *RAC*-encoded polyprotein (1438 residues) from accession PI 89772. Conserved features of the polyprotein are colored and identified as indicated in the above box.

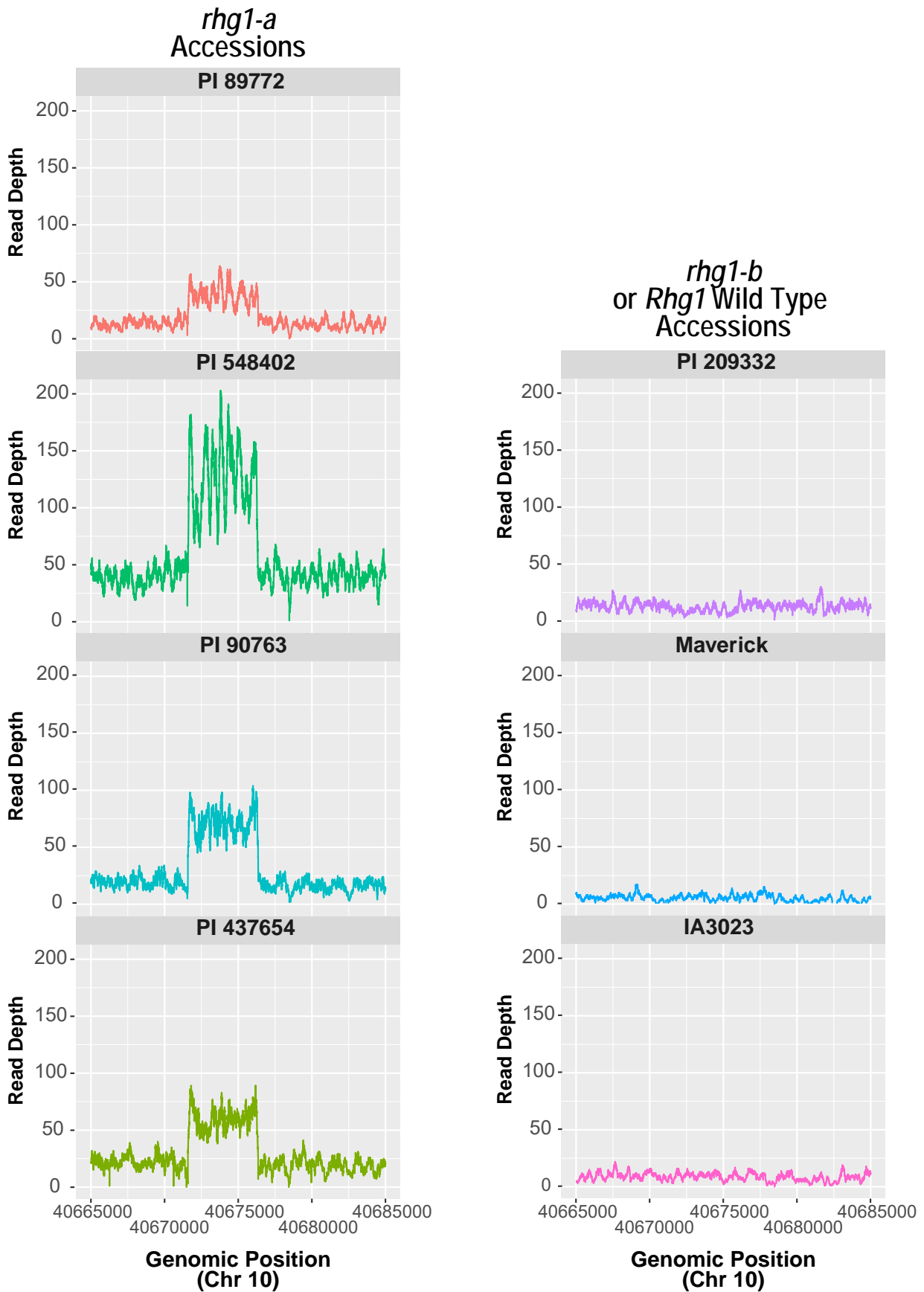


Fig. S4. Depth of short genomic reads that align to the *RAC* nucleotide sequence. The  $\alpha$ -SNAP-*RAC* insertion is not present within the Wm 82 soybean reference genome, thus reads align to the genomic position of the Chr 10 *RAC*-like element (99.7% nucleotide identity to *RAC*). Read depth for *rhg1-a* accessions (PI 89772, PI 548402, PI 90763, PI 437654) is ~3 to 4 fold greater than for the *rhg1-b* accessions (PI 209332, Maverick) and *Rhg1*<sub>WT</sub> accession (IA3023), which do not contain Chr 18 *Rhg1*  $\alpha$ -SNAP-*RAC* insertions.

**A**

```

RAC      AAGCACATGTTATAAGAAACATGGAATGCCCTCTAATCACGATACTAGAAACAAGAATAG
C10      AAGCACATGTTATAAGAAACATGGAATGCCCTCTAATCACGATACTAGAAACAAGAATAG
          *****

RAC      TGGAAGGAAAGCTTGCCTCATTGTGGCAAGATGGGACACACTGTGGACGTGTGTTATCG
C10      TGGAAGGAAAGCTTGCCTCATTGTGGCAAGATGGGGCACACTGTGGACGTGTGTTATCG
          *****

RAC      AAAACACGGGTACCCACCGGGTTACACGCCGGTTACAAGCCCTATGGAGGAAGAACCAC
C10      AAAACACGGGTACCCACCGGGTTACACGCCGGTTACAAGCCCTATGGAGGAAGAACCAC
          *****

```

**B**

SoySNP 50K	<i>rhg1-a</i>	<i>rhg1-b</i>
ss715629217	C	C
ss715629233	C	C
ss715629242	C	C
ss715629248	C	C
ss715629260	G	G
ss715629264	<b>A</b>	<b>G</b>
ss715629266	T	T
ss715629273	G	G
ss715629280	G	G
ss715629286	<b>T</b>	<b>C</b>
ss715629288	<b>C</b>	<b>T</b>
ss715629291	A	A
ss715629296	C	C
ss715629298	<b>G</b>	<b>A</b>

**C**

Accession	RAC SNP?	Detected <i>Rhg1</i> $\alpha$ -SNAP	<u>True</u> <i>rhg1-a</i> ?
PI 417441	No	HC $\alpha$ -SNAP	No
PI 507148	No	WT $\alpha$ -SNAP	No
PI 458094	No	HC $\alpha$ -SNAP	No
PI 408304	No	HC $\alpha$ -SNAP	No
PI 603438B	No	HC $\alpha$ -SNAP	No
PI 567319B	Present	LC $\alpha$ -SNAP	Yes
PI 567234B	Present	LC $\alpha$ -SNAP	Yes

Fig. S5. Sequence of the *RAC* SNP (ss715606985), nucleotide SNP signatures of consensus *rhg1-a* (*Rhg1* low-copy) and *rhg1-b* (*Rhg1* high-copy) haplotypes, and detected *Rhg1*  $\alpha$ -SNAP transcripts from *RAC*<sup>+</sup> or *RAC*<sup>-</sup> accessions with consensus *rhg1-a* SNP-signatures. (A) Short DNA alignment of *RAC* and the Chr10 (“C10”) element showing the position of the ss715606985 SNP (G to A) associated with *RAC* presence. (B) Consensus SNP signatures for *rhg1-a* and *rhg1-b* haplotypes, as identified by Lee *et al*, 2015. (C) *Rhg1*  $\alpha$ -SNAP alleles detected from genomic DNA subclones of *RAC*<sup>+</sup> or *RAC*<sup>-</sup> accessions which have a consensus *rhg1-a* SNP signature. HC, LC or WT refers to high-copy (*rhg1-b*), low-copy (*rhg1-a*) or wild-type *Rhg1*  $\alpha$ -SNAP alleles, respectively.

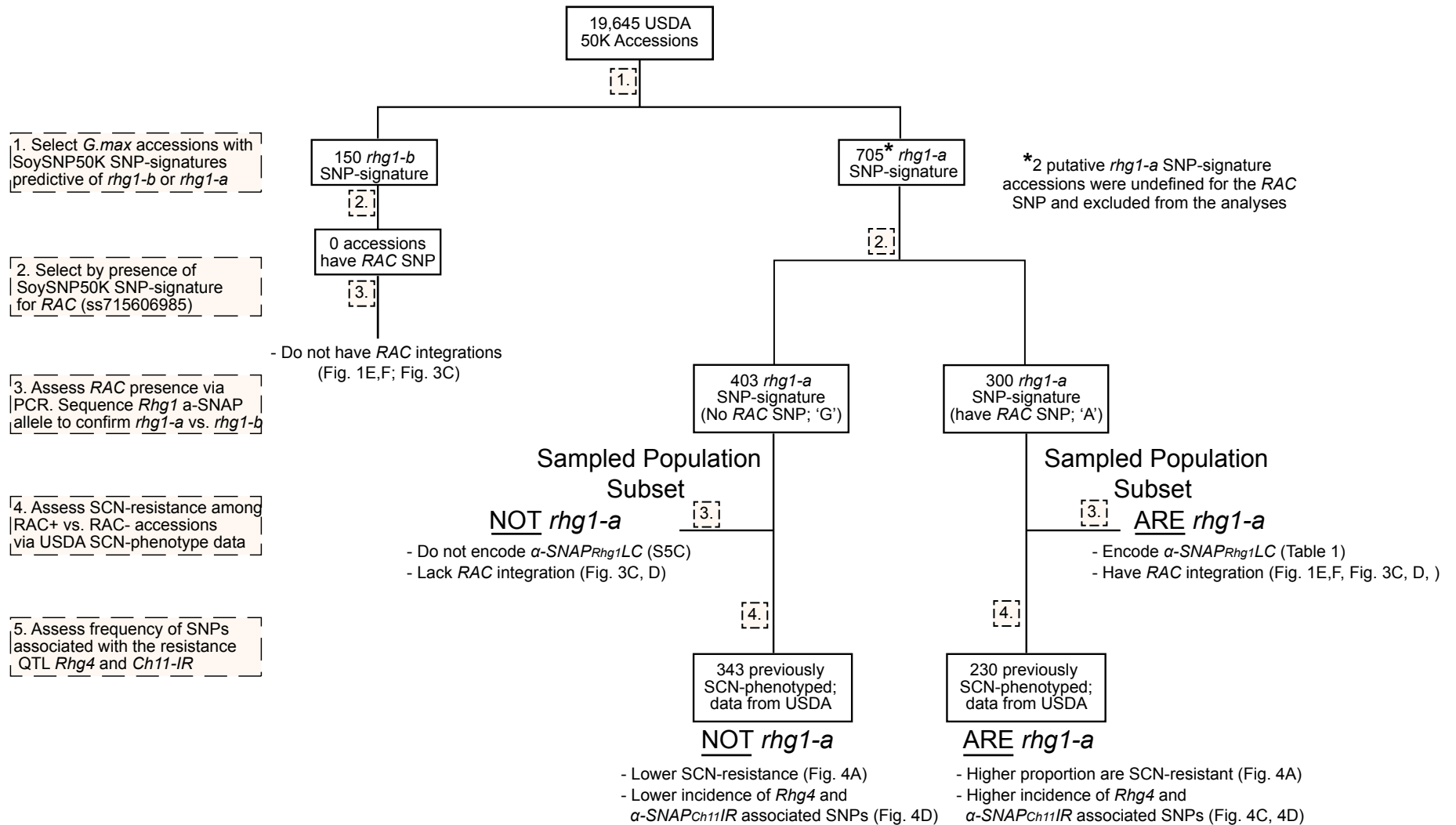
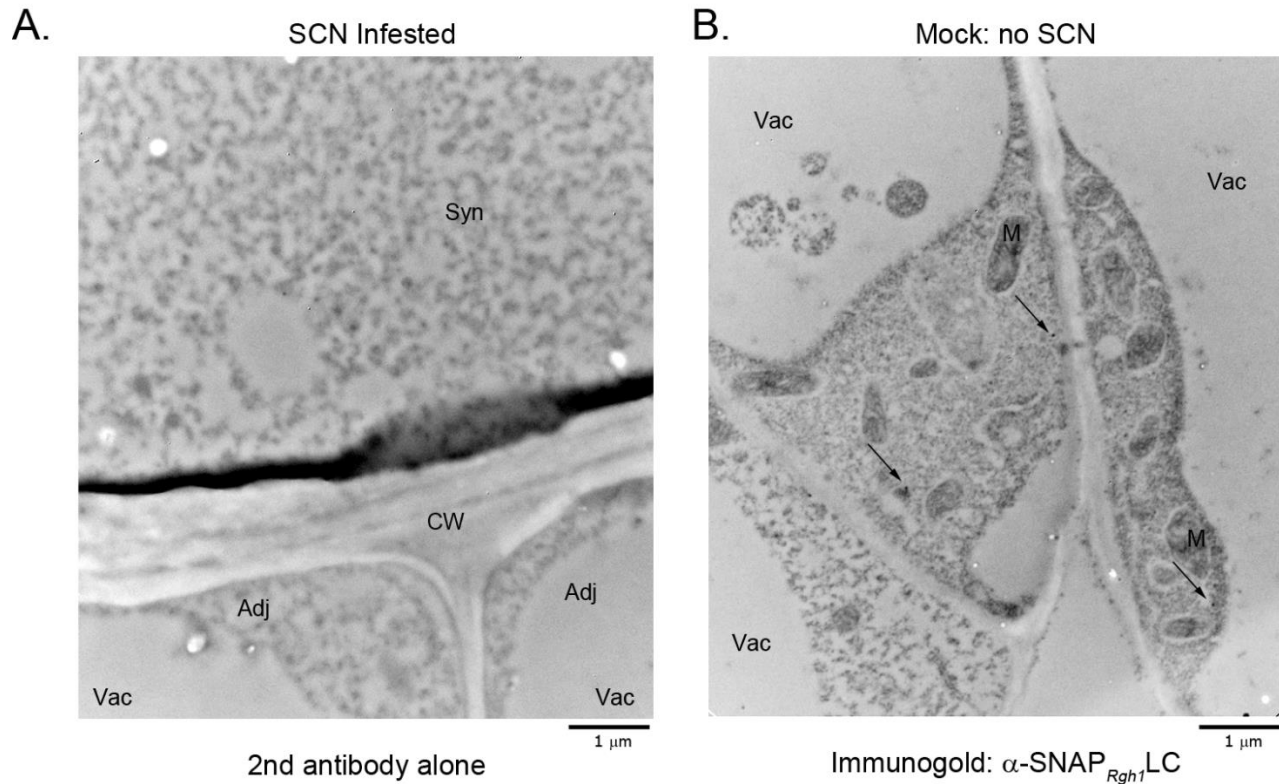


Fig. S6. Flow chart summarizing findings regarding  $RAC^+$  vs.  $RAC^-$  accessions, which otherwise have consensus SNP signatures for *rhg1-a*. The RAC SNP is useful to identify true *rhg1-a* accessions with strong SCN-resistance. See manuscript text for full description.





**Fig. S7**  $\alpha$ -SNAP<sub>Rhg1</sub>LC immunolabeling in 'Forrest' roots is highly specific. (A) Representative electron microscope image showing immunogold labeling using only secondary goat anti-rabbit antibody on SCN-infested 'Forrest' roots. Without prior incubation with the anti- $\alpha$ -SNAP<sub>Rhg1</sub>LC antibody, few gold particles were observed in syncytial vs. adjacent non-infected cells from the same root sections mounted on different grids. Syn, syncytium cells; Adj, adjacent cells. (B) Electron micrograph of a mock-inoculated 'Forrest' root after immunogold label detection using the anti- $\alpha$ -SNAP<sub>Rhg1</sub>LC primary antibody. Arrows indicate three immunogold particles across three root cell regions.

(In both A and B, CW, cell wall; M, mitochondrion; Vac, vacuole; bar = 1  $\mu$ m.)

Accession	<i>Rhg1</i> Copy	SCN Resistant	<i>Rhg1</i> WT $\alpha$ -SNAP	<i>Rhg1</i> HC $\alpha$ -SNAP	<i>Rhg1</i> LC $\alpha$ -SNAP	WT Junction	RAC Junctions
Williams 82	1	—	✓	—	—	✓	—
'Forrest'	3	✓	—	—	✓	—	✓
PI 90763	3	✓	—	—	✓	—	✓
PI 437654	3	✓	—	—	✓	—	✓
PI 89772	3	✓	—	—	✓	—	✓
PI 548316	7	✓	✓	✓	—	✓	—
PI 88788	9	✓	✓	✓	—	✓	—
PI 209332	10	✓	✓	✓	—	✓	—

Table S1. Summary of features of *Rhg1* haplotypes in different soybean accessions, using data from previously published work and the current study.

## Table S2

Location	BLAST Score	RAC Identity	Note
Ch10	8540	99.70%	Copia ORF fully intact; would encode 1438 residue polyprotein
Ch18	8076	97.60%	Integrated within intron 1 of <i>Glyma.18G268000</i> (anti-sense, copia ORF interrupted at residue 993)
Ch09	6522	90%	Copia ORF adjacent and intact; anti-sense to <i>Glyma.09G206300</i> (ABC transporter)
Ch14	4731	82%	copia ORF interrupted at residue 545
Ch02	4666	82%	copia ORF interrupted at residue 115
Ch20	4644	82%	Intronic integration within <i>Glyma.20G250200</i> (BAR domain)
Ch04	4080		
Ch01	3927		
Ch15	3910		
Ch05	3479		
Ch06	3285		
Ch03	3057		
Ch07	3036		
Ch16	2947		
Ch13	2922		
Ch08	2156		
Ch12	2040		
Ch17	1563		
Ch11	827		
Ch19	553		
<i>P. vulgaris</i> Ch02	378		

Table S2. *RAC*-like elements identified from NBLAST searches of *RAC* against the Williams 82 reference genome at Phytozome.org. The highest *RAC*-like match identified in the genome of the common bean (*Phaseolus vulgaris*) is also included. Only a single element (with highest identity to *RAC*) is shown for each of the n = 20 chromosomes of soybean. Several *RAC*-subfamily elements within the soybean reference genome are also inserted into host genes, and/or include elements with intact ORFs.