2008

The Multigeneic Rhg1 Locus: A Model For The Effects on Root Development, Nematode Resistance and Recombination Suppression

David A. Lightfoot
Southern Illinois University Carbondale, ga4082@siu.edu

Ali Srour

Jawad Afzal

Navinder Saini

Published in *Nature Precedings* http://hdl.handle.net/10101/npre.2008.2726.1(2008)

Classification: Biological Sciences; Agricultural Sciences

The Multigeneic *Rhg1* Locus: A Model For The Effects on Root Development, Nematode Resistance and Recombination Suppression.

Ahmed J. Afzal[1], Ali Srour[1], Navinder Saini[1,2], and David. A. Lightfoot[1,3]

1 Department of Molecular Biology, Microbiology and Biochemistry, Southern Illinois University at Carbondale, IL, 62901, USA.

2 Biotechnology Center, University of Jabalapur, India

3 Corresponding author; Genomics Core Facility; Department of Plant Soil and Agricultural Systems, and Center for Excellence in Soybean Research, Teaching and Outreach, Southern Illinois University at Carbondale, Carbondale, IL 62901-4415, USA. Tel 618 453 1797; Fax 618 453 7457; ga4082@siu.edu

Keywords **Segregation; Multigeneic Locus; *Rhg1;* Soybean; Resistance; Cyst Nematode**

- **Manuscript information:** 44 (7) text pages (including 3 Tables and 5 figures).
- **Abbreviations:** G10; gene 10, laccase: G11; gene 11, ion antiporter; RLK; receptor like kinase.
- **Data depositions:** TMD1 marker, bankit1155666, FJ520231; RLK at rhg1 gene bankit1154506; SIUC-Satt122, bankit1155667; 73P06 complete sequence bankit########.

Abstract

Soybean (*Glycine max* L. Merr.) resistance to any population (HgType) of *Heterodera glycines* I., the soybean cyst nematode (SCN), requires a functional allele at *rhg1*. An apoptosis-like response in the giant cells formed at the nematode feeding site results about 24-48 h after feeding commences. The response may be mediated by the 3 genes within the *rhg1* locus, a receptor like kinase (RLK), a laccase and an ion anti-porter . This study aimed to identify the role of the genes. Used were near isogeneic lines that contrasted at their *rhg1* alleles(NILs). Five features of the *rhg1* locus, the candidate genes and their nascent proteins were elucidated. First, evidence for a syntenic gene cluster on Lg B1 was found. Second, the effectiveness of SNP probes for distinguishing homeolog sequence variants on LgB1 from alleles at the *rhg1* locus on LgG was shown. Third, analysis of polymorphism among heterozygotes found NIL 34-33 segregating at the *rhg1* locus showed that the resistant allele was dominant in NILs and segregated in phase with a modifier. Fourth, the total absence of recombination events among the NILs between the RLK and other 2 genes eliminated the possibility of a monogeneic *rhg1* locus. Cosegregation of an unlinked locus was detected and a mechanism for segregation distortion inferred. Evidence for the presence transcripts and proteins encoded by the three genes at the *rhg1* locus was shown. Finally, an effect on root development was discovered. A model for multigeneic resistance based on developmental control of root growth is presented.

\body

Introduction

Soybean (*Glycine max* L. Merr.) seed yield losses due to root infestation by *Heterodera glycines* I. soybean cyst nematode (SCN) have been severe [1]. Losses have occurred since the crop was domesticated, since SCN has been the most widespread and damaging soybean pathogen worldwide. Soybean resistance to SCN was found in about 1% of pre-domesticated and early domesticated Plant Introductions (PI) [2, 3]. The development of partially resistant cultivars by gene introgression and other disease management measures have limited the costs of SCN infestation to the US soybean producers to about $1 billion/year in seed yield losses.

Soybean cyst nematode, like many plant parasitic nematodes, lives as an obligate endoparasite of plant roots that can use many alternate hosts [4, 5] . Over the past 50 years in the US, the number of Hg Types (ex. races) of SCN that are recognized has expanded from 4 in the 1960's to 16–20 of a possible 1,024 to date [6, 7]. Directed breeding for cultivar resistance will continue to select for new Hg types, so new resistance genes and new alleles will be continually needed. Resistant cultivars have a Female Index (FI) of less than 10% of the cyst numbers on susceptible cultivars in parallel tests [2]. Though arbitrary, the 10% measure of resistance often approximates to the economic loss thresh-hold in US soils (about 25 cyst/ 100 cm^3 of soil). Genetic diversity is found both in the field and in the commonly used inbred cyst populations like PA3, Hg type 0. PA3 was derived from a field population by incomplete inbreeding. Therefore, recombination due to sexual reproduction, transposon derived genome plasticity and/or mutation may continue to generate diversity in this population and other populations of the same Hg Type [8].

Variation among the host plant roots response to SCN has been associated with light, temperature and genetic purity such that genetic identities and environmental conditions must be rigorously controlled during host pathogen assays [2,3]. In both resistant and susceptible cultivars, the interaction between the nematode and soybean root passes through several discernable phases [9,10, 11]. However, resistance or susceptibility of the soybean is not induced until females establish a feeding site. The feeding site develops to provide a giant cell with its own secondary root-type vasculature providing nutrients to the female cyst. Cell to cell contact occurs at the syncitia through a stylet sufficiently narrow to prevent the passage of proteins and other molecules greater than 20 kD. Membrane to membrane contact was inferred. Secretions from several glands of SCN contain plant growth regulators and other bioactive factors.

Inheritance of resistance to SCN was first reported in the PI ‘Peking’ , and three recessive gene symbols (*rhg1–rhg3*) were assigned to the underlying loci [6]. One locus, *rhg1*, provides the major portion of resistance to SCN Hg type 0 (race 3) and Hg type 1.3.5.6.7.8, (race 14) across many genotypes whether they were derived from Peking, ‘PI437654’, ‘PI88788’, ‘PI209332’ or ‘PI90763’ [12] . However, given the evolution of the soybean nematode interaction is ancient and complex [4] it is likely that the locus contains several genes with each contributing partly to the activity of the locus [13].

The cytological studies suggests the *rhg1* driven Peking-type resistances share mechanisms of giant cell breakdown (pronounced necrosis and cell wall appositions) not seen in PI88788 type resistances in response to Hg type [9]. The differences in the mechanism of giant cell breakdown in Peking and PI88788 may derive from distinct alleles at *rhg1* and/or other defense-associated loci [2,14]. The *rhg1* locus was repeatedly located to a sub-telomeric region of the soybean molecular linkage group G by many studies [12-20]. However, some mapped resistance sources have an *rhg1*-like locus (required for resistance to all races) at another location in SCN resistant PIs, including Lg B1 [21], mid LgG, [22] and LgB2 [23]. Therefore, functional paralogs of *rhg1* may exist among the duplicated regions of the soybean genome [24].

Genes underlying resistance to Hg type 0 (PA3, race 3) have been mapped with greatest accuracy using recombinant inbred lines (RILs) and near isogenic lines (NILs) derived from the cross of 'Essex' by 'Forrest' [13, 24-27]. Forrest provided a unique set of tools for genomics. Forrest introgressed only resistance to Hg type 0 from Peking [28] that also resists two other Hg types. Only *rhg1* and *Rhg4* were introgressed into Forrest [19] . Several NIL populations segregating for *rhg1* and/or *Rhg4* were developed from the cross of Essex by Forrest [13, 29-31]. Genomic analysis identified three genes and their intergeneic regions within the 42 kbp identified as the locus [13]. Within the region the Forrest genes [13] showed with many allelic differences compared to susceptible genotypes 'A3244' [32] and 'Williams 82' (www.phytozome.net) with nine alleles recognized among PIs and four among resistant PIs.

The action of the *rhg1* resistance alleles has complex effects. First, *rhg1* alone was necessary, but not alone sufficient, for resistance to all known Hg-biotypes [12] . Second, some *rhg1* alleles restrict seed germination [33]. Third some interacting alleles are needed prevent zygote death are co-inherited (located on LgM as judged by RFLP markers) [15]. Fourth, some *rhg1* alleles inhibit seed yield at harvest in the absence of the disease [34-36]. Therefore, multiple gene or locus interactions were inferred that both underlie resistance and also alter plant development [13].

The three genes within the markers bounding the *rhg1* locus in Forrest included the RLK, a laccase and a predicted sodium/hydrogen antiporter [13]. Immediately outside the locus (on the basis of recombination events) were two predicted proteins of unknown function. Only the RLK and laccase and the 46.1 Kd predicted proteins were present in EST collections derived from roots. These three genes and their integeneic regions may interact to provide resistance to SCN.

Here, a molecular basis for resistance to SCN is inferred from five features of the *rhg1* locus, the RLK candidate gene and its nascent protein.

Results and Discussion

The soybean genome is hypothesized to be the product of a diploidized tetraploid. Therefore a detailed molecular analysis of the *rhg1* locus required that paralogs and syntenic gene clusters be identified. Probes developed from BAC 73P06 (Figure 1) were used.

Syntenic paralogs of the *rhg1 locus*

To identify *rhg1* paralogs, primers specific to the conserved regions of *rhg1* leucine-rich repeat and *rhg1*- kinase domain were designed. The conserved regions were determined by aligning DNA sequences from known *rhg1*-like genes in different soybean cultivars and other plant species. PCR-amplified products were radiolabeled and used as probes against the Forrest BIBAC libraries. The hybridizing BAC colonies were confirmed by Southern hybridizations to purified DNA. There were five positive clones for the LRR probe (Supplemental Figure 1) and three were also positive from the kinase hybridizations. One of the identified clones (B21d09) contained the *rhg1*, found on scaffold 121 whereas the other three clones contained RLK paralogs (B10a18, B55i16 and H38f23). B10a18 and B55i16 were on Lg A1 scaffold 15; H38f23 was on Lg B1 and scaffold 139 (69,100-144,000). Significantly, this BAC and scaffold 139 contained a complete set of syntenic genes for a second *rhg1* locus(the RLK, laccase both antiporters, the kinase and the helicase; Figure 1). The DNA markers sequence paralogs were present but more diverged except TMD1 that was highly conserved. Sequence analysis of the alleles in Williams 82, ‘Asgrow’ 3244 and Forrest RLK at *rhg1* showed 99% amino sequence identity. Among the paralogs DNA sequence identity was

high (~92% in geneic region; Supplemental Figure 2). Amino acid identity was 84% in the LRR, 86% in the transmembrane domain and 94% in the kinase domain. The laccase and the antiporter also showed 85-96% amino acid identity.

Both of the RLK paralogs were located by BLAT of microsatellite markers and BES to sequence scaffolds and were in regions where loci with functions similar to *rhg1* were located Lg A1 (the RLK) [21] and LG B1 (the syntenic cluster) [23]. The paralogs may encode proteins that recognize novel race biotypes or substitute for *rhg1* following activation in certain PIs or crosses [21-23].

Allele discrimination

Since paralogs with homeolog sequence variants (HSVs) appeared to exist for each gene in the cluster it was important to distinguish alleles precisely and separately from HSVs. In the NIL population RLK alleles were distinguished using a SNPs from the LRR region (Supplemental Figure 3); and the SIUC-TMD1 marker (Figure 2). In the intergeneic region between the RLK and laccase the SNP probe 10893 was used (Figure 1). Within the laccase was used an indel in the first intron, G10 probe from exon 2-3 and SNP probe 37583 in exon 6. In the anti-porter probe G11 and SNP probe 375821 was used. Each probe could detect polymorphism among the alleles of the three genes at *rhg1* in the ExF derived NILs but not at the paralogous loci.

Dominant, recessive or co-dominant nature of *rhg1*

In NIL 34-33 segregating at the *rhg1* locus, 4 plants were heterozygous at the TMD1 (Figure 2; Table 1) and Sac5 markers (results not shown). The frequency of heterozygousity among F5:13 generation seed was surprisingly high [13,31]. The existence of these plants suggests that fixation is selected against or heterozygosity is selected for, at this locus in these and related NILs.

The cyst scores, for all plants in the NIL population corresponded with the respective alleles at the *rhg1* locus. For the four heterozygous plants, polymorphic at TMD1, the cyst score correspond to those for resistant plants. Therefore, the *rhg1* locus was dominant in this set of NILs infested with this HgType.

Both recessive and co dominant roles have been assigned for the *rhg1* locus. In both past and recent studies with PIs, resistance encoded by *rhg1* was reported as recessive [2,14] whereas previously in NILs, the *rhg1* locus was reported to be co-dominant [19] but without single plant to marker allele associations. Co-dominant and recessive roles of plant disease resistance loci are rare and unusual [37, 38]. However, on the basis of the segregation pattern at the intrageneic TMD1 (intron) and Sca5 (promoter) markers, in NIL 34-33 background, the *rhg1* locus was shown to be dominant (Figure 2; Table 2). The discrepancies in dominance among different populations may be associated with the genetic background the gene resides in and may result from interactions among genes at the *rhg1* locus and/or modifier genes at other loci [15]. The phenotype at the rice blast resistant locus, *Xa3* is also influenced by the genetic background. The gene at the locus behaves differently in different genetic backgrounds, even displaying dominance reversal in one case [39].

Inhibition of root growth by alleles of *rhg1* in the NILs

When counting the cysts with prior knowledge of the allele at *rhg1* it was noted that root mass and vigor appeared to differ among genotypes. Measurements of root mass showed a significant difference among NILs that were associated with the allele at *rhg1* or linked loci (Figure 3; Table 3). Across several experiments, both NILs that were pure breeding susceptible and NILs that segregated some susceptible lines had higher root masses than their SCN resistant counterparts. This phenomenon might underlie the global association of resistance to SCN with low seed germination, seedling vigor, stand formation and ultimately seed yield [33,35].

The recombination events found among the six Hg Type 0 susceptible PIs [13] suggests that the action of *rhg1* requires elements to the distal side of the RLK intron, possibly one or all of the 3 polymorphisms found in the intracellular kinase of the complete RLK. Some mutations in the kinases of other plant RLKs are known to be lethal [40,41]. Kinase mutants can be lethal in many cases [42, 43]. Therefore, it may be the kinase at the *rhg1* locus that underlies restricted root growth in resistant genotypes directly or after some sort of interaction.

The three genes at *rhg1* are expressed in both resistant and susceptible soybean roots

The presence of the RLK at *rhg1* mRNA and protein in roots was confirmed by RT-PCR and Western hybridization (Figure 4). The *rhg1* transcript was detected under both inoculated and non-inoculated conditions in the both the resistant cultivar Forrest and resistant NIL34-23 and the susceptible cultivar Essex and susceptible NIL34-3.

The quantitative PCR used to determine differences in transcript abundance between infected and uninfected cultivars showed the mRNA was increased about 2 fold following SCN inoculation.

Expression judged by examination of EST libraries in silico, cDNA libraries by hybridization and mRNA populations by RT-PCR showed the RLK (Figure 4), laccase [44] and 46.1 Kd hypothetical transporter protein (unpublished) were transcribed in both non-infested roots and SCN-infested roots. However, paralogs were detected for each gene as judged by multiple amplicons from cDNA with laccase probes (G10) [44] and antiporter probes (G11; not shown). Therefore, genes in the syntenic *rhg1* paralog locus might influence *rhg1* activity by cooperation or competition.

Evidence for segregation distortion at *rhg1* from the absence of recombination events

Using the complete set of microsatellite and SNP probes across *rhg1* no recombination events have been found between TMD1 and SIUC Satt75 in the resistant haplotype within the region encompassing the 3 genes at *rhg1*. Used have been the ExF RIL population (n = 100) [19]; the set of SCN resistant PIs (n=112) [13]; NILs with recombination events between Satt309 and Satt214 collected from two populations ExF 34 (2,000) and ExF11 (2,000) [31]; and RILs with recombination events between Satt309 and Satt038RILs collected from the RxH population (n=975) and FxH population (n = 725) [45]. Here, the region from the RLK to the Na H antiporter was analyzed all available markers [13, 42] polymorphic in ExF (Figure 1). Again no recombination events were found. An absence of recombination events in a region can

have one of several causes among them; lack of homology; regional inversions; condensed heterochromatin; and recombinant allele lethality. The first three are not occurring at the *rhg1* locus since DNA sequences are collinear over 100 kbp, even though the locus was an introgression from a PI [13]. Therefore, recombinant allele lethality is the most likely cause.

A hypothetical model for recombinant allele lethality was developed. The kinase domain of the RLK at *Rhg1* was proposed as the killer element and the hypothetical protein was proposed as the target locus kept in the resistant state when the RLK is conferring SCN resistance (Figure 5). The laccase [44] trapped between these two genes and the intergenic regions are held in phase by the locus. It is possible the lethal nature of the resistant linkat is not fully suppressed (leaky) and results in the inhibition of root growth observed (Figure 3).

Conclusions

Suppression of recombination at *rhg1* was shown to center on the three genes at the core of the locus. Whether this causes the unusually high frequency of heterozygous plants or whether the helicase closely linked to the locus [13, 31] has local effects remains to be determined. A suppressor locus acting on *rhg1* was identified earlier [15]. The zygote or embryo lethal gene on Lg M that is co-inherited with *rhg1* [15] proved to have a homeolog that was near Satt594 on Lg G in ExF RILs (Table 1). Mapping the locus to a LG was difficult due to segregation distortion among resistant lines, but considering only susceptible lines carrying the susceptibility allele at *rhg1* the suppressor appears to be in the middle of LgG both in the map and in the sequence

scaffolds. The locus was fixed to the R haplotype in the NILs. Here we propose the locus be named suppressor of *rhg1* or *sup-rhg1.*

The discovery of a syntenic paralog to the gene cluster at *rhg1*raised significant barriers to reverse genetic approaches to the unequivocal proof that the RLK at *rhg1* candidate gene underlies part of the resistance to Hg type 0. Another barrier was the role of *rhg1* in normal plant development that can be inferred from the restricted root growth of NILs (Figure 3), some mutants in this gene (K. Meksem unpublished data) and the shoot effects measured in grafting experiments (A.J. Afzal published data). A third barrier to reverse genetics was the nematodes ability to inhibit RNAi activity (Dr. Chris Taylor unpublished data). Proof of *rhg1* function may require knock-outs of each of the paralogs (K. Meksem et al. unpublished data), or stable transformation to a new location (D. Simmonds and D.A. Lightfoot unpublished data), followed by measurements of genetic segregation. In each case, the analysis will be complicated by the co-dominant nature of the resistance gene in certain backgrounds [4, 20]. In fact, the possibility that the susceptible Essex allele, *rhg1*, promotes the establishment of parasitism by SCN must be explored.

The data presented here suggests the genes linked to the primary candidate RLK may encode factors involved in the modulation of *rhg1* activity. Possible roles include contributions to additive resistance; contributions to resistance in other resistance types (eg PI88788 and Toyohazu; R types 2 and 3) or contributions to the resistance to other Hg types [46]. Unlikely, in view of the susceptibility of segregation events within the interval from the RLK to Satt309 in both PI evolution and NIL segregation, is the hypothesis that the linked genes are factors necessary in

susceptible genotypes for SCN parasitism. The dominance of *rhg1* in NIL segregation also suggests the genes are active in resistant types and inactive in susceptible genotypes. In conclusion, the *rhg1* locus was inferred to be a complex of three genes assembled and co-inherited over long periods of selection for resistance to an pandemic pest, root parasitic nematodes [47].

The probes developed provide a high throughput alternative to satellite markers for marker assisted selection, three allelic discrimination tools were developed for Taqman primer probes (Supplemental Figure 3). The first Taqman probe (1040) could successfully discriminate resistant types 1 and 2 from susceptible haplotypes 2, 3 and 4. Marker 506 could distinguish R types 2 and 3 from other haplotypes. Marker 2050 distinguished among susceptible types. These tools will facilitate molecular breeding.

Materials and Methods

Plant materials

Many of the genetic materials were described previously [13, 29, 31]. Briefly, the seeds of RILs and the NIL populations derived from the cross of Essex by Forrest were obtained from Dr. Paul Gibson at Southern Illinois University at Carbondale in 1995 and were increased from 1995 to present at the Agronomy Research Center [48]. NIL populations were developed and maintained as described in [31]. Genetic identity and purity were checked after increase and before each experiment with 5–10 SSR markers/ line and DNA from 5–10 seeds/line. All lines are available on request as seed. Seeds of NIL 34-23 (resistant haplotype between markers Satt 214 to Satt 570) and NIL 34-3 (susceptible haplotype from the most telomeric marker Satt 214 to the Sat122-Satt 570 interval) were obtained at the F5:13 generation. Genotypes were

rhg1rhg1Rhg4Rhg4 for NIL 34-3 and *Rhg1Rhg1Rhg4Rhg4* for NIL 34-23 whereas NIL34-33 contained both those and *Rhg1rhg1Rhg4rhg4* in different plants.

Near isogenic line populations

Seed of soybean were obtained from the seed store at SIUC managed by Dr. Lightfoot. Seed of NIL 34-33 (polymorphic haplotype between markers Satt 214 to Satt 570) was obtained at the F5:13 generation. The three genotypes found within NIL 34-33 were; *rhg1Rhg1Rhg4Rhg4; rhg1rhg1Rhg4Rhg4* and *Rhg1Rhg1Rhg4Rhg4.*

SCN inoculations

Soybean plants were grown in 5 l buckets, each containing 20 cones in a randomized setup. Each bucket contained a 1:1 ratio of sand soil mix. The containers were placed in a water bath in the SIUC greenhouse. Growth conditions were a 14h light cycle, day time temperature of 30°C and a nighttime temperature of 22°C. The humidity was maintained at approximately 40-50% (v/v). Infection with Hg Type 0 SCN populations consisted of inoculating 2,000 eggs to each 4 day old seedling. Inoculated soybean plants were removed from the cones; 30 days post inoculation and cyst numbers counted.

The NIL experiments used single-plant replications. The cultivars ‘Lee 74’, ‘Essex’ and ‘Hutcheson’ were used as susceptible controls (Niblack et al., 2003). The differentials or indicator lines and the associated female indices (FI) were ‘PI54840’ (FI 7%), PI 88788 (FI 2%), PI90763 (FI 1%), PI437654 (FI 0%), ‘PI 209332’ (FI 1%), ‘PI89772’ (FI 2%) ‘PI548316’ (FI 8%) and ‘PI548402’ (FI 3%). Therefore, the standard

differentials showed this HG Type to be 0 [7] (Niblack et al. 2003) corresponding to race 3 [49].

DNA and RNA extraction

DNA was isolated following [50] modified as follows. Briefly, 100 mg of frozen plant tissue was ground, 600 µl preheated (65°C) extraction buffer, incubated at 65°C for 1 hour, cooled and centrifuged at 10,000 g for 15 minutes. The supernatant was decanted and 5 µl of RNase (5 mg/ml) was added at 37°C for one hour. The aqueous phase was extracted first by the addition of equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) v/v, incubated for 2 min by slow inversion followed by centrifugation at 10,000 g for 15 minutes and secondly by the addition of equal volume of chloroform: Isoamyl alcohol (24:1). DNA was collected by the addition of iso-propanol to the supernatant, followed by centrifugation at 12,000 g for 5 min. The DNA pellets were washed twice with 70% (v/v) ethanol, dried and finally dissolved in 30 µl Tris buffer. Concentrations of DNAs were calculated by measuring absorbance at 260 and 280 nm.

Total RNA was isolated with Trizol (Invitrogen Cat. No. 15596-026, Carlsbad, CA, USA), according to the manufacturer's instructions. The RNA pellet was dissolved in 40µl of DEPC treated MQ water and quantified by measuring absorbance at 260 nm. First strand cDNA was synthesis carried out using oligo dT primers using a cDNA synthesis kit, according to manufacturer (Invitrogen). Presence of the *rhg1* mRNA was confirmed by PCR analysis using a *rhg1* intron flanking primers pair: *rhg1*-int-F-LRR (ATT TGA ATC AGA AGT CAG TGT) and *rhg1*-int-R-LRR (TCT GGT CTA ATC TCT TCC AGC; Supplemental Table 2).

NIL genotyping by microsatellite markers linked to *rhg1*

About 50 ng of DNA was used for microsatellite analysis on PAGE after [35]. The microsatellite markers from the RLK, SIUC-TMD1 and SIUC-Sac 5, were used to genotype resistant and susceptible segregants from NIL 34-33 [13] (Supplemental Table 2). Amplification reactions for BARC-Satt markers for NIL analysis on agarose gels were performed after [51].

Taqman assays of alleles the RLK within the rhg1 locus

The SNP genotyping assay within the gene encoding the RLK was performed using a custom Taqman^{TM} Kit. Three probes were designed 1486, 506 and 2040 to distinguish the 8 commonest alleles of the RLK (Supplemental Table 2). Only probe 1486 was polymorphic in Essex Forrest and the derived NILs. A 242 bp amplification reaction was carried out using an *rhg1* LRR forward primer: 5' CAG AGA ACA ACC TCC TTG 3' and an *rhg1* LRR reverse primer: 5' CAG AAC CTG AGA GGC TAT 3'; IDT DNA, Coralville, IA, USA) with the following discriminatory probe pair. Probe 1:5'-Fam-TAT TCC TTC AAG CAT TGC AAA CAT TTC CTC G-BHQ1-3' and Probe 2: 5' Hex -TAT TCC TTC AAG TAT TGC AAA CAT TTC CTC GC-BHQ1-3'. Primer and probe optimization were done by using different combinations of each pair and optimizing to optimal signal strength and balanced fluorophore intensity. The PCR reaction was carried out using a 3 step PCR protocol with one hold at 95°C for 10 minutes followed by 35 cycles that included a denaturation cycle of 95°C for 30 sec, annealing at 58°C for 10 seconds and an extension at 68°C for 20 sec.

SNP assays

Primers for SNPs within the *rhg1* locus on Lg G were used in fine melt curve assays using the ABI7900 with HTM software as described previously [52] with the following modifications. Briefly, genomic DNA was used; multiple amplicon sizes were detected on PAGE gels; melt curve data were normalized by both local (local background value was subtracted from the intensity value of sample) and global metrics. Three SNP primers used were as described in [53]. They were; AX196295 10893 at 54,040 bp between the laccase and RLK; AX196295 37583 CR-G at 62,107 in the laccase; and AX196295 37581 CR-G at 64,929 bp in the hypothetical gene (Supplemental data). Additional primers were designed to detect Essex to Forrest polymorphisms among the three genes and intergeneic regions within the *rhg1* locus.

Southern hybridization

Southern hybridizations were performed following the standard procedure described in [54]. Total genomic DNA was digested with restriction enzymes, separated by electrophoresis on a 0.8% (w/v) agarose gel and transferred onto a positively charged nylon membrane. After hybridization, the corresponding bands were visualized by exposure of X-ray film for 24–48 h.

Total root protein extraction

Protein from root material was isolated from infested and non-infested roots of Forrest and Essex after [55, 56]. Briefly, 2g of the finely ground frozen root was

resuspended in 5mL of Tris buffered phenol (pH 8.8) and 5mL of extraction buffer. The solution was vortexed vigorously and centrifuged at 5,000g for 15 minutes. After removal of the top phase (phenol) the bottom phase was back extracted with Tris buffered phenol (5ml) and an equal volume of the extraction buffer. Proteins were pelleted by centrifugation at 20,000g for 20 minutes and washed. The pellet was dried and resuspended in SDS loading buffer. Total protein concentration was determined using a non-interfering protein assay [57].

SDS–PAGE and Western hybridization

SDS–PAGE of total plant proteins from Essex and Forrest followed by Western hybridization was carried out according to [58] with the following modifications. For the Western hybridizations, a custom made antibody generated against a peptide CTL SRL KTL DIS NNA LNG NLP ATL SNL S from the LRR domain of RHG1 was used (Alpha diagnostics, San Antonio, Texas). As a secondary antibody, an anti rabbit IgG HRP was used (GE healthcare, Milwaukee, Wisconsin).

Acknowledgements

The physical map was based upon work supported by the National Science Foundation under Grant No. 9872635. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. The continued support of SIUC, College of Agriculture and Office of the Vice Chancellor for Research to DAL

and AS is appreciated. The authors thank Dr. Q. Tao and Dr, H.B. Zhang for assistance with fingerprinting. We thank Dr. C. Town and Dr. C. Foo at TIGR for the BES. We thank Jeffry Shultz for assistance with probing the BAC libraries. We thank the DOE CJSP for release of the WGS reads and scaffolds. We thank the Government of India for support of the Fellowship for NS.

References

1. Wrather JA, Anderson TR, Arsyad DM, Tan Y, Ploper LD, Porta-Puglia A, Ram HH, Yorinori JT (2001) Soybean disease loss estimates for the top ten soybean-producing countries in 1998. *Can J Plant Path* 23:115-121
2. Arelli PR (1994) Inheritance of resistance to *Heterodera glycines* race 3 in soybean accessions. *Plant Dis* 78:898–900
3. Arelli PR, Sleper DA, Yue P, Wilcox JA (2000) Soybean reaction to Races 1 and 2 of *Heterodera glycines. Crop Sci* 40:824-826
4. Hussey RS, Grundler FMW (1998) Nematode parasitism of plants. In: Perry RN Wright DJ (eds), " The physiology andbiochemistry of free-living and plant parasitic nematodes." CAB International Wallingford, USA, pp 213–243
5. Keen NT Roberts PA (1998) Plant parasitic nematodes: Digesting a page from the microbe book. *Proc Natl Acad Sci* USA 95:4789-90
6. Caldwell BE, Brim CA, Ross JP (1960) Inheritance of resistance of soybeans to the cyst nematode. *Heterodera glycines. Agron J* 52:635–636
7. Niblack TL, Noel GR, Lambert KL (2003) The Illinois SCN type test: practical application of the Hg type classification system. *J Nematol* 35:355–345
8. Bekal S, Niblack TL, Lambert KN (2003) A chorismate mutase from the soybean cyst nematode *Heterodera glycines* shows polymorphisms that correlate with virulence. *Mol Plant Microb Interact* 16:439-46
9. Mahalingam R, Skorupska HT (1996) Cytological expression of early response to infection by *Heterodera glycines* Ichinohe in resistant PI 437654 soybean. *Genome* 39:986-998
10. Davis EL, Hussey RS, Baum TJ (2004) Getting to the roots of parasitism by nematodes. *Trends Parasitol* 20:134-41
11. Gao BR, Allen EL, Davis TJ, Baum RS, Hussey (2004) Molecular characterization and developmental expression of a cellulose-binding protein gene in the soybean cyst nematode *Heterodera glycines. Int J Parasitol*

34:1377-83
12. Concibido VC , Diers BW, Arelli PR (2004) A decade of QTL mapping for cyst nematode resistance in soybean. *Crop Sci* 44:1121-1131
13. Ruben E, Jamai A, Afzal J, Njiti VN, Triwitayakorn K, Iqbal MJ, Yaegashi S, Bashir R, Kazi S, Arelli PR, Town CD, Ishihara H, Meksem K, Lightfoot DA (2006) Genomic analysis of the *rhg1* locus: candidate genes that underlie soybean resistance to the cyst nematode. *Mol Genet Genom* 276:503-516
14. Brucker E, Niblack T, Kopisch-Obuch FJ, Diers BW (2005) The effect of *rhg1* on reproduction of *Heterodera glycines* in the field and greenhouse and associated effects on agronomic traits. *Crop Sci* 45:1721-1727
15. Webb DM, Baltazar BM, Arelli PR, Schupp J, Clayton K, Keim P, Beavis WD (1995) Genetic-mapping of soybean cyst-nematode race-3 resistance loci in the soybean PI 437654. *Theor Appl Genet* 91:574-581
16. Chang SJC, Doubler TW, Kilo VY, AbuThredeih J, Prabhu R, Freire V, Suttner RJ, Klein III JH, Schmidt ME, Gibson PT, Lightfoot DA (1997) Association of loci underlying field resistance to soybean sudden death syndrome (SDS) and cyst nematode (SCN) race 3. *Crop Sci* 37:965-971
17. Concibido VC, Lange DA, Denny RL, Orf JH, Young ND (1997) Genome mapping of soybean cyst nematode resistance genes in 'Peking' PI 90763 and PI 88788 using DNA markers Crop Science 37:258-264
18. Cregan PB, Mudge J, Fickus EW, Marek LF, Danesh D, Denny D, Shoemaker RC, Matthews BF, Jarvik T, Young ND (1999) Targeted isolation of simple sequence repeat markers through the use of bacterial artificial chromosomes. *Theor Appl Genet* 98:919-928
19. Meksem K, Pantazopoulos P, Njiti VN, Hyten DL , Arelli PR, Lightfoot DA (2001) 'Forrest' resistance to the soybean cyst nematode is bigenic: saturation mapping of the *Rhg1* and *Rhg4* loci. *Theor Appl Genet* 103:710-717
20. Meksem K, Iqbal MJ, Ruben E, Hyten DL, Triwitayakorn K, Lightfoot DA (2001) Conversion of AFLP bands into high-throughput DNA markers. *Mol Genet Genom* 265:207-14
21. Vierling RA, Faghihi J, Ferris VR, Ferris JM (1996) Association of RFLP markers with loci conferring broad-based resistance to the soybean cyst nematode (*Heterodera glycines*). *Theor Appl Genet* 92:83-86
22. Wang D, Arelli PR, Shoemaker RC, Diers BW (2001) Loci underlying resistance to Race 3 of soybean cyst nematode in *Glycine soja* plant introduction 468916 *Theor Appl Genet* 103:561-566
23. Yue P, Arelli PR, Sleper DA (2001) Molecular characterization of resistance to *Heterodera glycines* in soybean PI438489B. *Theor Appl Genet* 102:921-928
24. Shultz JL Kurunam D, Shopinski K, Iqbal MJ, Kazi S, Zobrist K, Bashir R, Yaegashi S, Lavu N, Afzal AJ, Yesudas CR, Kassem MA, Wu C Zhang HB,

Town CD, Meksem K, Lightfoot DA (2006) The Soybean Genome Database (SoyGD): A browser for display of duplicated polyploid regions and sequence tagged sites on the integrated physical and genetic maps of *Glycine max. Nucl Acids Res* 34:D758-65

25. Lightfoot DA (2008) Soybean genomics: Developments through the use of cultivar Forrest. *Internat J of Plant Genom* 2008:1-22
26. Meksem K, Zobrist K, Ruben E, Hyten DL, Quanzhou T, Zhang HB, Lightfoot DA (2000) Two large-insert soybean genomic libraries constructed in a binary vector: applications in chromosome walking and genome wide physical mapping. *Theor Appl Genet* 101:747-755
27. Wu C, Sun S, Nimmakayala P, Santos FA, Meksem K, Springman R, Ding K, Lightfoot DA, Zhang HB (2004) A BAC- and BIBAC-based physical map of the soybean genome. *Genom Res* 14:319-26
28. Hartwig EE, Epps JM (1973) Registration of Forrest soybeans. Crop Sci13:287
29. Njiti VN, Doubler TW, Suttner RJ, Gray LE, Gibson PT, Lightfoot DA (1998) Resistance to soybean sudden death syndrome and root colonization by *Fusarium solani* f sp *glycine* in near-isogenic lines. *Crop Sci* 38:472-477
30. Meksem K, Doubler TW, Chancharoenchai K, Njiti VN, Chang SJC, Arelli PR, Lightfoot DA (1999) Clustering among loci underlying soybean resistance to *Fusarium solani* SDS and SCN in near-isogenic lines. *Theor Appl Genet* 99:1131-1142
31. Triwitayakorn K, Njiti VN, Iqbal MJ, Yaegashi S, Town CD, Lightfoot DA (2005) Genomic analysis of a region encompassing *QRfs1* and *QRfs2*: genes that underlie soybean resistance to sudden death syndrome. *Genome* 48:125-138
32. Hauge BM, Wang ML, Parsons JD, Parnell LD (2001) Nucleic acid molecular and other molecules associated with soybean cyst nematode resistance. *WO 01/51627 PCT/US01/00552 Patent # 20030005491*
33. Kopisch-Obuch FJ, McBroom KL, Diers BW (2005) Association between SCN resistance loci and yield in soybean. *Crop Sci* 45:956-965
34. Mudge J, Cannon SB, Kalo P, Oldroyd GE, Roe BA, Town CD, Young ND (2005) Highly syntenic regions in the genomes of soybean *Medicago truncatula* and *Arabidopsis thaliana. BMC Plant Biol* 5:15
35. Yuan J, Njiti VN, Meksem K, Iqbal MJ, Triwitayakorn K, Kassem MA, Davis GT, Schmidt ME, Lightfoot DA (2002) quantitative trait loci in two soybean recombinant inbred line populations segregating for yield and disease resistance. *Crop Sci* 42:271-277
36. Kassem MA Shultz J, Meksem K, Cho Y, Wood AJ, Iqbal MJ, Lightfoot DA (2006) An updated 'Essex' by 'Forrest' linkage map and first composite interval map of QTL underlying six soybean traits. Theor Appl Genet 113:1015-1026
37. Li CW, Bai YL, Jacobsen E, Visser R, Lindhout P, Bonnema G (2006) Tomato defense to the powdery mildew fungus: differences in expression of genes in

susceptible monogenic- and polygenic resistance responses are mainly in timing. *Plant Molec Biol* 62:127-140

38. Nieto C, Piron F, Dalmais M, Marco CF, Moriones E, Gómez-Guillamón ML, Truniger V, Gómez P, Garcia-Mas J, Aranda MA, Bendahmane A (2007) EcoTILLING for the identification of allelic variants of melon eIF4E a factor that controls virus susceptibility. *BMC Plant Biol* 7:34-44
39. Xiang Y, Cao Y, Xu C, Li X, Wang S (2006) Xa3 conferring resistance for rice bacterial blight and encoding a receptor kinase-like protein is the same as *Xa26. Theor Appl Genet* 113:1347-1355
40. Lease KA, Lau NY, Schuster RA, Torii KU, Walker JC (2001) Receptor serine/threonine protein kinases in signalling: analysis of the erecta receptor-like kinase of *Arabidopsis thaliana. New Phytol* 151: 133-143
41. Wang X, Goshe MB, Soderblom EJ, Phinney BS, Kuchar JA, Li J, Asami T, Yoshida S, Huber SC, Clouse SD (2005) Identification and functional analysis of in vivo phosphorylation sites of the *Arabidopsis* BRASSINOSTEROID-INSENSITIVE1 receptor kinase. *Plant Cell* 17 1685–1703
42. Sessa G, D'Ascenzo M, Martin GB (2000) Thr38 and Ser198 are *Pto* auto-phosphorylation sites required for the *AvrPto–Pto*-mediated hypersensitive response. *EMBO J* 19: 2257-2269
43. Kollewe C, Mackensen AC, Neumann D, Knop J, Cao P, Li S, Wesche H, Martin MU (2004) Sequential auto-phosphorylation steps in the interleukin-1 receptor-associated kinase-1 regulate its availability as an adapter in interleukin-1 signaling. *J Biol Chem* 279: 5227-5236
44. Ahsan R, Iqbal MJ, Afzal AJ, Jamai A, Meksem K, El Shemy H, Lightfoot DA (2008) Analysis of the activity of the soybean laccase encoded within the *Rfs2/rhg1* locus. *Curr Iss Mol Biol* (in press)
45. Prabhu, RR, Njiti V, Johnson JE, Schmidt ME, Klein RJ, Lightfoot DA (1999) Selecting soybean cultivars for dual resistance to cyst nematode Sudden Death Syndrome with two DNA markers. *Crop Sci* 39:982-987.
46. Nelsen N, Zhigang L, Warner AW, Matthews BF, Knap HT (2003) Genomic polymorphisms identifies a subtilisin-like protease near the *Rhg4* locus in soybean. *Crop Sci* 44:265–273
47. Afzal AJ, Wood AJ, Lightfoot DA (2008) Plant receptor-like serine threonine kinases: Roles in signaling and plant defense. *Mol Plant Microb Inter* 21:507-517
48. Lightfoot DA, Njiti VN, Gibson PT, Kassem MA, Iqbal MJ, Meksem K (2005) Registration of the Essex X Forrest recombinant inbred line mapping population. *Crop Sci* 45:1678-1681
49. Riggs RD, Schmitt DP (1988) Complete characterization of the race scheme for *Heterodera glycines. J Nematol* 20: 392-395
50. Saghai-Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance chromosomal location and population dynamics. *Proc Natl Acad Sci USA* 81: 8014-8018

51. Shultz JL, Kazi S, Bashir R, Afzal JA, Lightfoot DA (2007) The development of BAC-end sequence-based microsatellite markers and placement in the physical and genetic maps of soybean. *Theor Appl Genet* 114:1081-90
52. YuanJ, Haroon M, Lightfoot DA, Pelletier Y, Liu Q, Li X-Q (2008) A High-Resolution Melting Approach for Analyzing Allelic Expression Dynamics. *Curr Issue Mol Biol* (in press)
53. Hyten DL, Choi IY, Song Q, Shoemaker RC, Nelson RL, Costa JM, et al. Highly variable patterns of linkage disequilibrium in multiple soybean populations. *Genetics* (2007) 175:1937–1944.
54. Meksem K, Doubler TW, Chancharoenchai K, Njiti VN, Chang SJC, Arelli PR, Lightfoot DAl (1999) Clustering among loci underlying soybean resistance to *Fusarium solani* SDS and SCN in near-isogenic lines. *Theor Appl Genet* 99:1131-1142
55. Hurkman WJ, Tanaka CR (1986) Solubilization of plant membrane proteins for analysis by two-dimensional gel electrophoresis. *Plant Physiol* 81:802-806
56. Hajduch M, Ganapathy A, Stein JW, Thelen JJ (2005) A systematic proteomic study of seed filling in soybean: Establishment of high-resolution two-dimensional reference maps expression profiles and an interactive proteome database. *Plant Physiol* 137:1397-419
57. Sheffield J, Taylor N, Fauquet C, Chen SX (2006) The cassava (*Manihot esculenta* Crantz) root proteome: Protein identification and differential expression. *Proteomics* 6:1588-1598
58. Afzal, AJ, Lightfoot DA (2007) Inclusion bodies contain RHG1 folding intermediates: A novel refolding protocol for protein purification. *Protein Expr Purific* 53: 346–355.

Table Legends

Table 1: DNA marker segregation within the SCN resistant lines in the ExF34 derived NIL population lines 1-40 showing some resistance to HgType 0 the phenotype was associated only with segregation near Satt309. Female Index was the mean of 5 plants repeated once (10 plants total). A represents the Essex allele; B the Forrest allele and H identifies lines that are heterogenous and may contain heterozygous plants (eg. NIL34-33).

Table 2: Cyst score and phenotype for plants from Figure 2. The associations between phenotypic scores and marker scores are perfect for all plants. Plant 18 (*) and 19 were switched (**).

Table 3: Association of root growth in NILs with root development in seedlings at 28 days after germination with SCN infestation.

Figure Legends

Figure 1: Marker map of the genomic region around *rhg1* and the homeolg of *rhg1* with locus ideograms. Marker anchors are shown for each. Scaffold 139 was anchored by Satt484 at 1,005,915 bp. The B73P06 left BES was homeologous to scaffold 139 at 121,470 bp and homologous to scaffold 121 at 1,675,980 bp. The right BES was at homologous to scaffold 121 at 1,759,500bp (83,520 bp) and weakly homeologous to scaffold 139 at 44,800 bp (76,670 bp). The H38F23 left BES was homeologous to scaffold 139 at 69,150 bp and homologous to scaffold 121 at 1,725,800 bp. The right

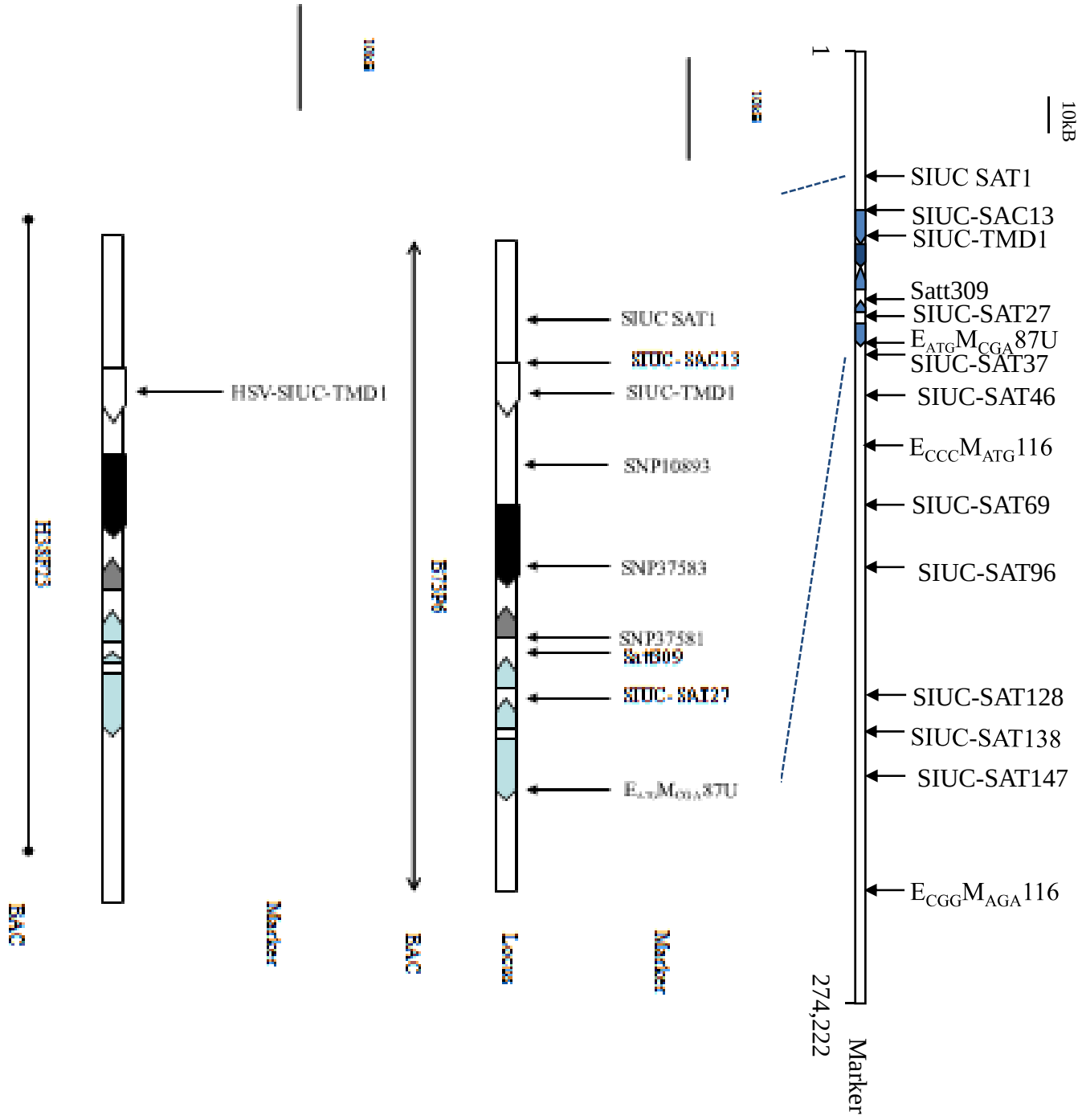
BES was at homologous to scaffold 139 at 144,000 bp (74,850 bp) and weakly homeologous to scaffold 121 at 1,657,130 bp (68,670 bp). The marker Satt TMD1 (gi:56718383) has a homeolog of 858 bp at 79,000bp on scaffold 139 and a homolog of 762 bp on at 1,714,500 bp on scaffold 121 but the primers are specific to amplify only the 762± 3 bp amplicon. The marker TMD-indel amplified a fragment from both homeologs (Supplemental Figure 3) of 303± 15 bp and 362 bp.

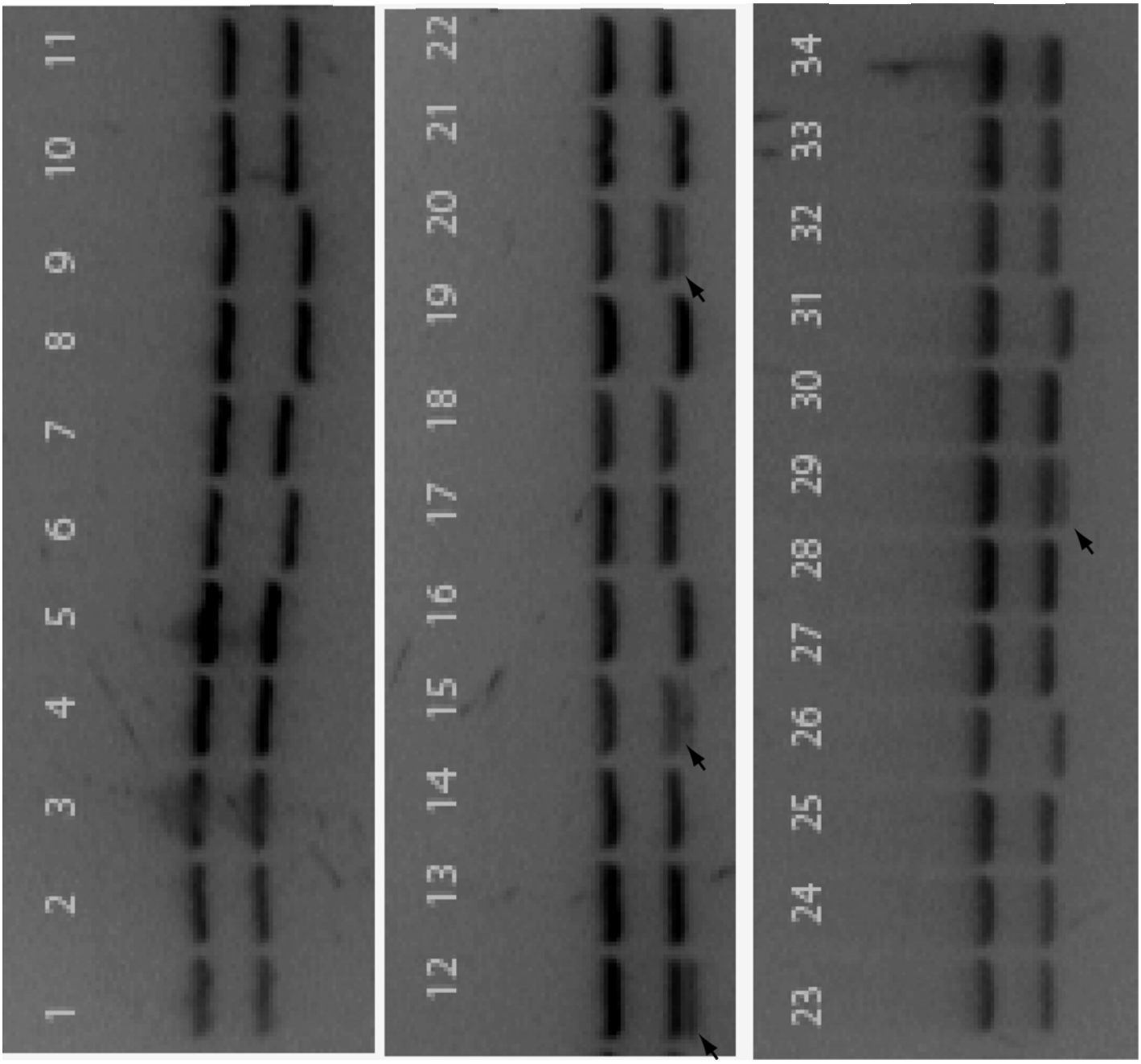
Figure 2: Analysis of NIL 34-33 with TMD1. Out of the 34 plants analyzed, 4 were heterozygous (Lanes 12, 14, 20 and 29). The four heterozygous lines had a resistant phenotype. SCN counts and gel scores for the plants are given in Table 3

Figure 3: The RHG1 protein alters root development. Panel A; soybean NILS at 2 weeks pre-SCN inoculation show different root morphologies, post inoculation root masses are not different (by 6 weeks). Therefore. *rhg1* inhibits germination and early root growth. Panel B; the root morphologies co-segregated with the allele in the RLK at the rhg1 locus as shown by the intragenic marker TMD1 (satellite in the intron).

Figure 4: Expression of the RLK at *rhg1* in soybean roots. Panel (A); Western hybridization using an anti-RHG1 antibody from; Forrest root (a); Essex root (b); expressed RHG1-LRR-Shrt (c); expressed RHG1-LRR-Long (d); and expressed RHG 4 (e). Panel (B); Agarose gel electrophoresis of cDNA amplified using *rhg1* LRR flanking primers from RIL 34-23; non-infested control (1); SCN infested RIL 34-23 (2); SCN infested RIL 34-3 (3); and Forrest genomic DNA (4). Negative control without template is shown in lane 5.

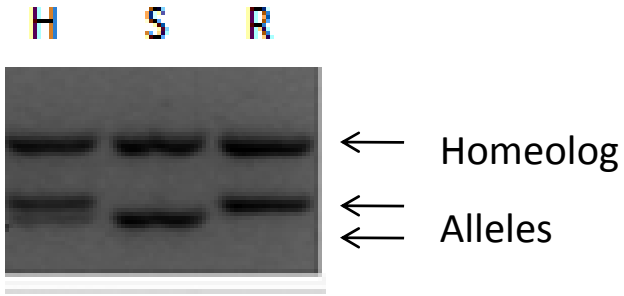
Figure 5: Model for the function of the *rhg1* locus in resistance to SCN. Black arrows show positive interactions, blue arrows show inhibitions. In this model four phenotypic events are controlled by the 3 genes at *rhg1* and four unlinked genes *Rhg2-4* and *Rzd1* (*sup-Rhg1*). The apoptosis caused in the giant cell of the nematodes feeding sites will occur in plant root cells in the absence of *Rzd1* allele in the coupled phase with the RLK. Root growth inhibition occurs despite suppression of the RLK.


10kB
1
274,222
Marker
SIUC SAT1
SIUC-SAC13
SIUC-TMD1
Satt309
SIUC-SAT27
$E_{ATG}M_{CGA}87U$
SIUC-SAT37
SIUC-SAT46
$E_{CCC}M_{ATG}116$
SIUC-SAT69
SIUC-SAT96
SIUC-SAT128
SIUC-SAT138
SIUC-SAT147
$E_{CGG}M_{AGA}116$
Marker
Locus
BAC
SIUC SAT1
SIUC-SAC13
SIUC-TMD1
SNP10893
SNP37583
SNP37581
Satt309
SIUC-SAT27
$E_{ATG}M_{CGA}87U$
BAC
Marker
HSV-SIUC-TMD1
BAC



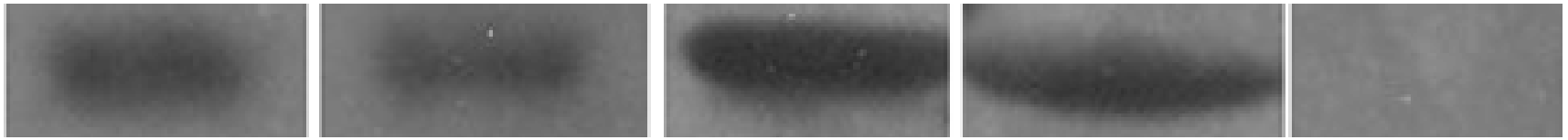
A.

B.



(A)

(a) (b) (c) (d) (e)



(B)

(1) (2) (3) (4) (5)