

Genetic and Physical Localization of the Soybean *Rpg1-b* Disease Resistance Gene Reveals a Complex Locus Containing Several Tightly Linked Families of NBS-LRR Genes

Tom Ashfield,¹ Anna Bocian,¹ Dan Held,¹ Adam D. Henk,¹ Laura Fredrick Marek,² Dariush Danesh,³ Silvia Peñuela,³ Khalid Meksem,⁴ David A. Lightfoot,⁴ Nevin D. Young,³ Randy C. Shoemaker,^{2,5} and Roger W. Innes¹

¹Department of Biology, Indiana University, Bloomington, IN 47405 U.S.A.; ²Department of Agronomy, Iowa State University, Ames IA 50010 U.S.A.; ³Department of Plant Pathology, University of Minnesota, St. Paul, MN 55108 U.S.A.; ⁴Department of Plant, Soil and General Agriculture, Southern Illinois University, Carbondale, IL 62901 U.S.A.; ⁵USDA-ARS, Corn Insect and Crop Genetics Research Unit, Iowa State University, Ames IA 50010 U.S.A.

Submitted 1 January 2003. Accepted 19 May 2003.

Alleles or tightly linked genes at the soybean (*Glycine max* L. Merr.) *Rpg1* locus confer resistance to strains of *Pseudomonas syringae* pv. *glycinea* that express the avirulence genes *avrB* or *avrRpm1*. We have previously mapped *Rpg1-b* (the gene specific for *avrB*) to a cluster of resistance genes (*R* genes) with diverse specificities in molecular linkage group F. Here, we describe the high-resolution physical and genetic mapping of *Rpg1-b* to a 0.16-cM interval encompassed by two overlapping BAC clones spanning approximately 270 kilobases. *Rpg1-b* is part of a complex locus containing numerous genes related to previously characterized coiled coil-nucleotide binding site-leucine rich repeat (CC-NBS-LRR)-type *R* genes that are spread throughout this region. Phylogenetic and Southern blot analyses group these genes into four distinct subgroups, some of which are conserved in the common bean, *Phaseolus vulgaris*, indicating that this *R* gene cluster may predate the divergence of *Phaseolus* and *Glycine*. Members from different subgroups are physically intermixed and display a high level of polymorphism between soybean cultivars, suggesting that this region is rearranging at a high frequency. At least five CC-NBS-LRR-type genes cosegregate with *Rpg1-b* in our large mapping populations.

Additional keyword: *RPM1*.

The outcome (disease or resistance) of many plant-pathogen interactions is governed by gene-for-gene-type resistance in which resistance genes (*R* genes) in the plant are matched by corresponding avirulence (*avr*) genes in the pathogen (Flor 1955). If either the *R* gene in the plant or the *avr* gene in the pathogen are absent, timely detection of the pathogen does not occur and disease ensues.

Numerous *R* genes involved in gene-for-gene-type resistance have been cloned from a variety of plant species. The

cloned *R* genes fall into several classes defined by shared structural motifs (Dangl and Jones 2001). The largest class is defined by the presence of a central nucleotide binding site (NBS) and C-terminal leucine rich repeats (LRRs). NBS-LRR *R* genes can be divided into two subfamilies differing in the N-terminal motifs present (Meyers et al. 1999; Pan et al. 2000). TIR-NBS-LRR *R* genes have motifs found in the toll/interleukin receptor proteins, while the CC-NBS-LRR genes (sometimes referred to as non-TIR NBS-LRR) contain a putative coiled-coil domain.

Genetic studies in several plant species have shown that resistance specificities are often clustered (Michelmore and Meyers 1998; Pryor and Ellis 1993). Such clustering is seen both for *R* genes specific for different races of the same pathogen (Hulbert and Bennetzen 1991) and for *R* genes conferring resistance to unrelated pathogens (Witsenboer et al. 1995). It is also common to find allelic series of *R* genes specific for different races of the same pathogen (Islam et al. 1989). These observations are reflected in the molecular architecture of *R* gene loci, which often consist of multigene families of linked sequences (Hulbert et al. 2001). Less commonly, *R* genes exist as single copy genes and are not linked to other NBS-LRR-type sequences. Such differences in locus organization would be predicted to have profound effects on the evolution of new specificities.

Occasionally, *R* genes with the same specificity are found in distantly related plant species. It is not known if this reflects the conservation of an ancestral specificity in more than one lineage or whether the specificity evolved independently more than once during the diversification of plant species (i.e., by convergent evolution). An example is provided by the recognition of the *Pseudomonas syringae* avirulence genes *avrB* and *avrRpm1* by Arabidopsis and soybean. The Arabidopsis *RPM1* gene has a dual specificity, conferring resistance to *P. syringae* strains expressing *avrB* or *avrRpm1* (Bisgrove et al. 1994). Conversely, soybean expresses *R* genes able to distinguish between these two avirulence genes. Some cultivars respond only to *avrB*, others only to *avrRpm1*. Genetic analysis has demonstrated that the corresponding *R* genes are either alleles or tightly linked genes at the *Rpg1* locus (Ashfield et al. 1995). The *Rpg1-b* gene is specific for *avrB*, *Rpg1-r* for *avrRpm1*.

Corresponding author: Roger W. Innes; Telephone: 1-812-855-2219; Fax: 1-812-855-6082; E-mail: rinnes@bio.indiana.edu.

The nucleotide sequence data reported are available in the DDBJ, EMBL, and GenBank databases under accession numbers AY278771 (RGA-39A), AY278772 (RGA-52A), and AY278773 (RGA-84A).

RPM1 has been cloned and has been found to be an *R* gene of the CC-NBS-LRR family (Grant et al. 1995). The *RPM1* gene is at a simple locus and is not clustered with other NBS-LRR genes. Interestingly, in all the naturally occurring *rpm1* lines examined, a deletion event has occurred at the *RPM1* locus, and the open reading frame is absent (Grant et al. 1995; Stahl et al. 1999).

We have previously mapped the *Rpg1* locus to a cluster of *R* genes in molecular linkage group F. Represented in the cluster are genes conferring resistance to a diverse selection of pathogens (Ashfield et al. 1998). *R* genes present include *Rpg1-b* and *Rpg1-r* (effective against *Pseudomonas syringae*) (Ashfield et al. 1998), *Rps3* (effective against *Phytophthora soja*) (Diers et al. 1992), *Rsv1* (effective against soybean mosaic virus) (Gore et al. 2002; Yu et al. 1994), and *Rpv1* (effective against peanut mottle virus) (Gore et al. 2002; Roane et al. 1983). All these genes map within 3 cM of the restriction fragment length polymorphism (RFLP) marker pR-45. Also present in the vicinity are QTLs conferring resistance to corn ear worm (Rector et al. 1999), two species of root-knot nematodes (Tamulonis et al. 1997a and b) and *Sclerotinia* stem rot (Arahana et al. 2001). Multiple alleles are found at several of the loci, increasing the number of specificities encoded by genes in this cluster still further (Ashfield et al. 1998).

Previous studies have shown that sequences related to NBS-LRR-type *R* genes are associated with this *R* gene cluster (Ashfield et al. 1998; Gore et al. 2002; Hayes and Saghai Maroof 2000; Hayes et al. 2000; Jeong et al. 2001; Peñuela et al. 2002; Yu et al. 1996). The majority of NBS-LRR genes in the region are of the CC-NBS-LRR-type, although at least one TIR-NBS-LRR sequence is also present (Hayes et al. 2000; Yu et al. 1996). Phylogenetic analysis reveals that the CC-NBS-LRR sequences in this cluster can be divided into several subgroups of highly related sequences (Jeong et al. 2001; Peñuela et al. 2002).

In this study, we present the high-resolution genetic and physical mapping of the *Rpg1-b* gene. We delimit *Rpg1-b* to a genetic interval of less than 0.2 cM and a physical interval encompassed by two overlapping BAC clones. We demonstrate that *Rpg1-b* resides in a complex locus containing numerous sequences related to previously characterized CC-NBS-LRR class *R* genes spread over several hundred kilobases. We demonstrate that these NBS-LRR genes belong to four distinct families of highly related sequences. Several of these NBS-LRR sequences cosegregate with *Rpg1-b*, and thus, represent strong candidates to encode this gene.

RESULTS

Construction of a physical map spanning *Rpg1-b*.

We have previously shown that the RFLP markers R45 and php2265 are tightly linked to *Rpg1-b*. In 149 recombinant inbred lines (RILs) from two populations, *Rpg1-b* cosegregated with R45 and was separated from php2265 in only a single line (Ashfield et al. 1998). Based on their tight linkage and presence in only single copy in the soybean genome these two markers were chosen to initiate construction of a physical map of the *Rpg1-b* region.

The R45 and php2265 clones were sequenced, and primers were developed. Both pairs of primers amplified only a single band from genomic DNA of soybean cultivar Williams 82 (*Rpg1-b* genotype). These primers were used to screen a bacterial artificial chromosome (BAC) library derived from Williams 82. This library consists of approximately 40,000 clones representing 4 to 5 genome equivalents (Marek and Shoemaker 1997). Three clones containing R45 and four clones containing php2265 were initially identified (Table 1). The BAC clones were ordered relative to each other by Southern blot hybridization with the probes derived from the ends of each BAC (Fig. 1B). The two sets of BAC clones did not overlap. To close this gap, we initiated a chromosome walk, which required us to orient the two BAC contigs relative to *Rpg1-b* and to increase our genetic resolution.

We had previously scored a small Flambeau (*rpg1-b Rpg1-r*) × Merit (*Rpg1-b rpg1-r*) RIL population for markers linked to the *Rpg1-b* locus. In all, eleven RIL families with recombination events between the *Rpg1-b* flanking markers HSP176 and B212 were identified (Ashfield et al. 1998). In this study, we further characterized these lines by rescoring them with polymerase chain reaction (PCR)-based markers derived from end sequences of the BACs described above (discussed below). Two of the lines (RI-12 and RI-61) had recombination events between markers EP-64A8r and php2265 (Fig. 1A). RI-61 was of particular interest because previous work had shown that it had inherited neither parental specificity, suggesting that it contains a recombination event that separates the *Rpg1-b* and *Rpg1-r* genes.

Additional informative recombinants in the *Rpg1* region were identified by screening two other RIL populations. The first set contained approximately 900 lines derived from a cross between cultivars Flyer (*Rpg1-b*) and Hartwig (*rpg1*), and the second set contained 94 lines derived from the cross between cultivars Essex (*rpg1-b*) and Forrest (*Rpg1-b*). The Flyer × Hartwig population was screened with markers R45,

Table 1. Summary of BAC library screens

Locus	Primers used to screen library ^a	Libraries screened ^b	BACs isolated (insert sizes)
R45	gtt caa gcc caa caa taa ca cca ccg aac cac gtt tag	IS	42I18 (130kb), 60I24 (130kb), 64A8 (130kb)
EP-42I18r	atc ctt ctg tat gga gaa ttg g aat ctg aca aaa gca ctt aca gat a	IS, SIU, UMN	14G17 (145kb)
EP-14G17r	gtt gtg act tca aca tcc cta tc caa aat ctt gta aat gta aac gtt	IS, SIU	52D1 (114kb), 21H16 ^c , 50N7 (95kb)
EP-52D1r	tgc aat tct ctt tac tgg ttg g tag agc gtg aaa gaa ttt gta gg	IS	21H16 ^c
EP-21F22f ^d	atc tgt gtg tag ggt ttt tga gt ctc aag agg aag aaa aat aat gg	IS, SIU, UMN	none
EP-95L8r	cac aaa atc gtc cac tac tca cat ggt caa tga aac tgt att	IS, UMN	21F22 (205kb), 18N17 (135kb), 162N9 (160kb)
php2265	cca acg gga gtt cag agt cc tga cct cag gga ttc cag tg	IS	12P6 (60kb), 39A5 (170kb), 84D9 (140kb), 95L8 (180kb)

^a Primer sequences read 5' to 3'.

^b SIU indicates the screening of both cv. Forrest libraries.

^c The size of the insert in BAC 21H16 could not be reliably determined.

^d An additional primer pair was used to screen the UMN library, tgc agg tga aga cac ttg ct; cag gta ttg cta ttg aag gtg ta.

EP-95L8r, and EP-84D9r (Fig. 1). Lines that displayed recombination events between one or more markers were then assayed for their *Rpg1-b* phenotype. These analyses revealed that *Rpg1-b* is located between markers R45 and EP-84D9r. The Essex (*rpg1-b*) × Forrest (*Rpg1-b*) population was scored with EP-64A8r and php2265. No additional recombinants were identified.

The nine informative recombinants from the Flambeau × Merit and Flyer × Hartwig populations were scored with the four markers defining the ends of the R45 and php2265 BAC contigs. This analysis enabled us to orient both contigs relative to *Rpg1-b*. BAC-end marker EP-42I18r from the R45 contig was found to cosegregate with *Rpg1-b*. BAC end EP-95L8r from the php2265 contig was separated from *Rpg1-b* and BAC end EP-42I18r in two of the recombinant lines (Fig. 1).

In an attempt to close the gap between the R45 and php2265 contigs, the EP-42I18r and EP-95L8r primers were used to re-screen the Williams 82 BAC library. No additional BAC clones were recovered with the EP-42I18 primers, and a single additional clone was recovered with EP-95L8r (Table 1). This clone, 21F22, overlapped extensively with BAC 95L8 and was found to contain marker php2265, indicating that it had been missed in the original screen with the php2265 primers. Analysis of restriction fragment patterns confirmed that the 21F22 BAC extended the php2265 contig. As expected, primers developed to the *Rpg1-b*-proximal end of BAC 21F22 failed to amplify from BAC 42I18, confirming that a gap still existed between the contigs. The 21F22 primers were developed into a PCR-based marker that was scored in the informative recombinants. This analysis revealed that BAC 21F22 had extended the

php2265 contig past an additional recombinational breakpoint (Fig. 1B).

Our inability to identify additional BAC clones with marker EP-42I18r indicated the Williams 82 library has a “hole” in this region. We therefore engaged in further rounds of screening, using three additional BAC libraries, two derived from cultivar Forrest and one from cultivar Faribault. Both cultivars Forrest and Faribault carry the *Rpg1-b* gene. Together, these three libraries represent approximately eleven genome equivalents.

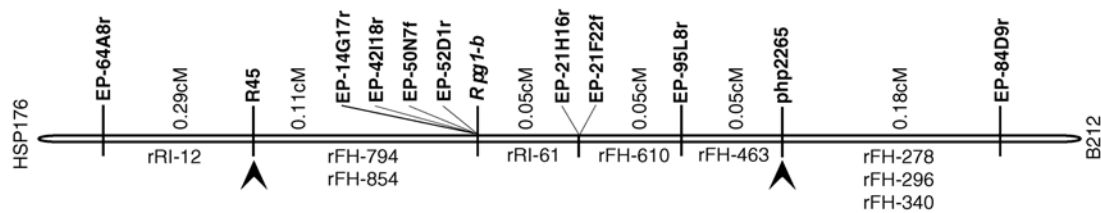
All three libraries were screened with the 42I18r primers. A single clone (BAC 14G17) was identified from the Faribault li-

Table 2. Sources of nucleotide binding site-leucine rich repeat (NBS-LRR) sequences

Sequence	Source	GenBank accession no.
Present within the R45 or php2265 BAC contigs ^a		
RGA-39A	Isolated from BAC IS_039_A05 using homology to RPM1	AY278771
RGA-52A	Identified from end sequencing of BAC IS_052_D01	AY278772
RGA-84A	Identified from end sequencing of BAC IS_084_D09	AY278773
NBSD-H8	Described by Peñuela et al. (2002)	AZ301636
Not present within the R45 or php2265 BAC contigs		
NBSD-H1	Described by Peñuela et al. (2002)	AZ301628
NBSD-H7	Described by Peñuela et al. (2002)	AZ301634
AF-F4	Described by Peñuela et al. (2002)	AZ301765

^a As defined by the presence of at least one band that hybridizes strongly to the NBS-LRR sequence under high-stringency conditions.

A Genetic Map



B Physical Map

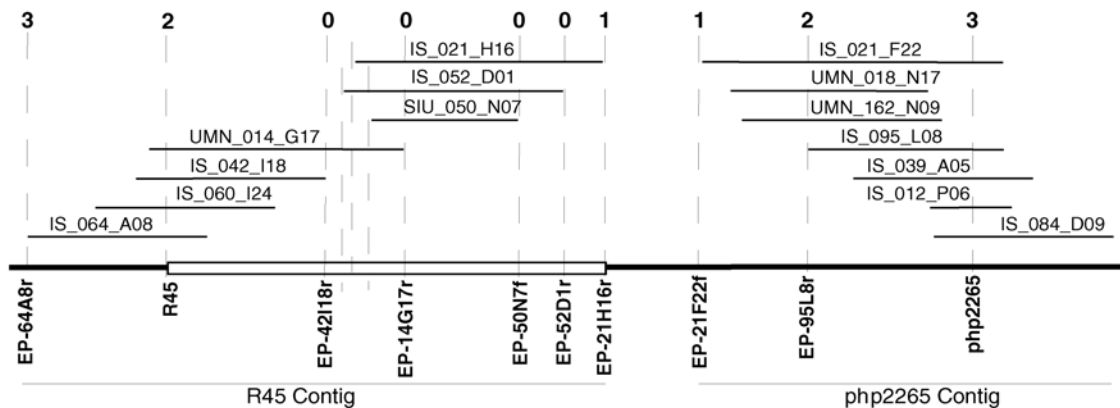


Fig. 1. High-resolution genetic and physical maps for the interval containing *Rpg1-b*. **A**, Fine-structure genetic map of the *Rpg1-b* region. *Rpg1-b* is located on soybean molecular linkage group F between markers HSP176 and B212 (Ashfield et al. 1998). Markers are shown in bold. Markers with the prefix EP are derived from BAC end sequences. Genetic distances between markers are indicated. Recombinant inbred lines (RILs) that contain recombination breakpoints between specific markers and have an r prefix. The two markers used to initiate BAC walks are indicated with arrows. This map represents the combined data obtained from three RIL populations: Flambeau × Merit, Flyer × Hartwig, and Essex × Forrest. Not all markers were polymorphic in all populations. **B**, BAC contigs linked to *Rpg1-b*. The contig on the left encompasses the genetically defined interval containing *Rpg1-b* (as indicated by the open box). polymerase chain reaction–confirmed overlaps between BACs are indicated by dashed lines. Numbers above dashed lines indicate the number of recombination breakpoints separating that locus from *Rpg1-b*. Marker names are given below each dashed line. The BACs and overlaps are not drawn to scale.

brary (Table 1). No clones were found in the Forrest libraries. PCR analysis using primers developed to the ends of the 14G17 insert confirmed that 14G17 had extended the R45 contig (Fig. 1B). The PCR marker corresponding to the *Rpg1-b* proximal end of 14G17 cosegregated with EP-421I8r and with *Rpg1-b* in the Flambeau × Merit informative recombinants (Fig. 1A). The 14G17r primers failed to amplify from BAC 21F22, indicating that the gap between the two contigs had not been closed.

All three libraries were also screened with primers developed from the *Rpg1-b* proximal end of BAC 21F22. No BACs were identified (Table 1), demonstrating that this region is not represented in any of the three libraries and suggesting that this region cannot be isolated as a BAC clone.

As noted above, the *Rpg1-b* proximal end of BAC 14G17 cosegregated with *Rpg1-b*. In an attempt to extend the contig beyond *Rpg1-b*, we rescreened the Williams 82 and Forrest BAC libraries with the 14G17r primers. Two BACs (52D1 and 21H16) were isolated from the Williams 82 library and one (50N7) from the *Hind*III Forrest library (Table 1). Primers were developed from all six BAC ends and were used to determine the arrangement of the new BACs relative to each other and to the existing R45 contig. This analysis demonstrated that, as expected, all three BACs extended the R45 contig with BAC 21H16 extending the furthest (Fig. 1B).

PCR-based markers were developed from the *Rpg1-b* proximal ends of BACs 21H16 and 52D1. These markers were scored in the informative recombinants. The 52D1r marker cosegregated with *Rpg1-b* in all nine recombinant lines. Significantly, the 21H16r marker cosegregated in all lines, except line RI61 in which the marker and *Rpg1-b* were separated by recombination, demonstrating that the R45 contig now extended beyond *Rpg1-b* (Fig. 1A and B). The 21H16r primers

failed to amplify from BAC 21F22, demonstrating that a gap still existed between the R45 and php2265 contigs.

We conclude that *Rpg1-b* has been localized to two overlapping BACs (14G17 and 21H16). As noted below, clone 21H16 is unstable and this prevented a definitive determination of physical distance. However, we estimate the two BACs to encompass an interval of approximately 270 kb, as deduced from restriction fragment analysis of these two BACs as well as overlapping BACs (Fig 1B).

Instability of BACs associated with the *Rpg1-b* locus.

As described above, the region immediately surrounding *Rpg1-b* is dramatically underrepresented in all four BAC libraries used in this study. We further observed that the *Rpg1-b* cosegregating BACs 21H16 and 52D1 inhibited the growth of their *E. coli* host strains and often accumulated deletions when grown in large-volume cultures for plasmid isolations (data not shown). These phenomena were mild with BAC 52D1 but very pronounced with BAC 21H16, suggesting that this region contains a DNA sequence that is toxic to *E. coli* even at low copy number.

Identification of NBS-LRR genes present within the BAC contig.

The majority of plant resistance genes cloned to date are of the NBS-LRR class, suggesting that *Rpg1-b* might also be of this type. We therefore attempted to identify NBS-LRR-type genes from the BAC contigs, to provide a source of *Rpg1-b* candidate genes for further analysis. Sequencing of BAC ends from the region initially identified two NBS-LRR genes, RGA-52A and RGA-84A (Table 2). An additional NBS-LRR gene (RGA-39A) was identified by hybridizing a probe derived from the cloned Arabidopsis *R* gene *RPM1* to blots of BACs from the region under reduced-stringency conditions.

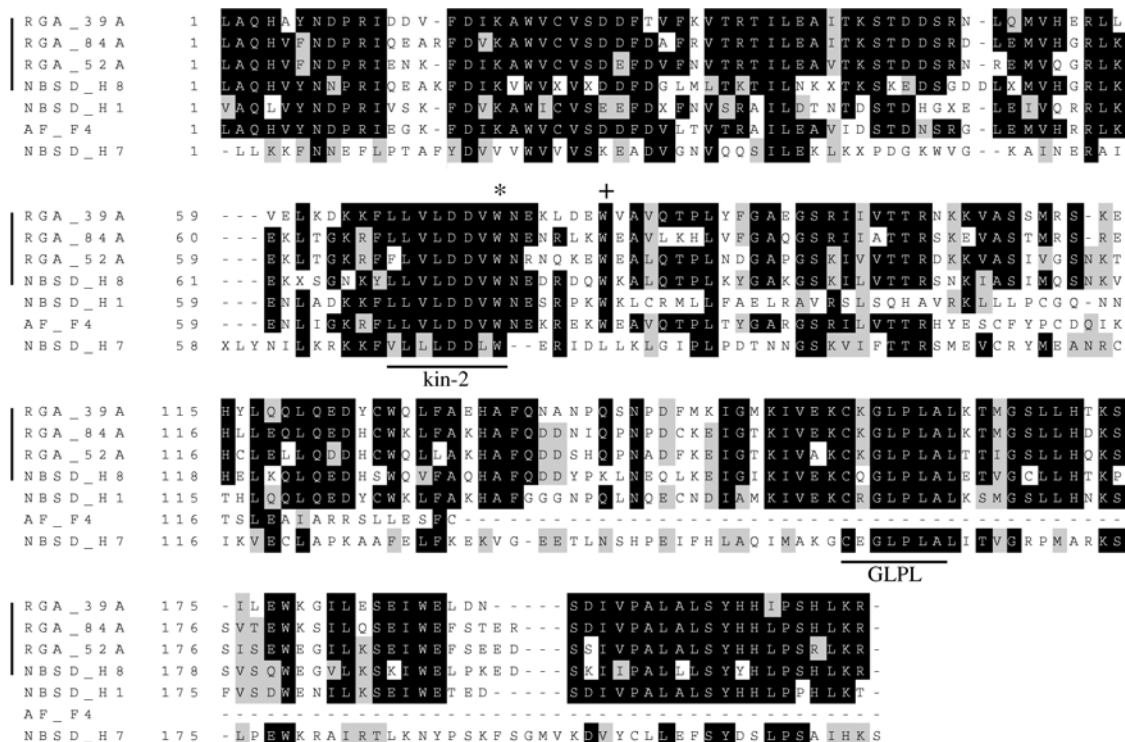


Fig. 2. Alignment of nucleotide binding site (NBS) regions from the NBS-leucine rich repeat (LRR) sequences used in this study. Amino acid sequences between the P-loop and RNBS-D motifs were aligned using Clustal X (Thompson et al. 1997). Sequences belonging to families represented within the R45 or php2265 BAC contigs are identified with a vertical line. The conserved kinase-2 and GLPL motifs are underlined. The conserved tryptophan found in the kinase-2 domain of most coiled-coil (CC)-NBS-LRR type *R* genes (Meyers et al. 1999) is marked with an asterisk. The tryptophan that defines a subgroup of CC-NBS-LRR sequences, including most of the NBS-LRR sequences mapped to the K644/B212 cluster of *R* genes (Jeong et al. 2001), is marked with a "+".

In addition to the three genes described above, four previously described NBS-LRR sequences were also included in our analysis. These sequences were isolated by Peñuela and associates (2002) using degenerate primers and PCR and have been mapped to the chromosomal region that contains the K644/B212 cluster of resistance specificities. To determine if any of these sequences were present within our BAC contigs, four clones (NBSD-H1, NBSD-H7, NBSD-H8, and AF-F4) representative of four subclasses were hybridized to our BAC contigs. As described later, only the NBSD-H8 clone detected highly related sequences within the contigs. Subsequent analysis revealed that NBSD-H8 shares 94% nucleotide identity (over a 431-nt overlap) with the *Rpg1-b* distal end of BAC 14G17, confirming that a NBS-LRR gene highly related to NBSD-H8 is present within the R45 contig.

The NBS regions of the clones described in Table 2 were aligned to more accurately determine the relationships between these genes. As expected, all contained the conserved kinase-2 and glycine-leucine-proline-leucine domains characteristic of NBS-LRR-type *R* genes (Fig. 2). Also present in all the sequences is a conserved tryptophan in the kinase-2 motif (Meyers et al. 1999) characteristic of the CC-NBS-LRR subfamily (Fig. 2). Consistent with this observation, the RGA-39A, RGA-52A, and RGA-84A sequences contain the variant of the RNBS-D domain (this motif is outside the aligned region) found in *R* genes of this type (Meyers et al. 1999). Six of the genes contain a second conserved tryptophan that has been used to define a subclass of CC-NBS-LRR genes, members of which have previously been mapped to the K644/B212 *R* gene cluster (Jeong et al. 2001).

Pairwise sequence comparisons between the four genes RGA-39A, RGA-52A, RGA-84A, and NBSD-H8, revealed that these sequences display between 72 and 83% nucleotide and 57 to 74% amino acid identity within the NBS region.

Phylogenetic analysis of NBS-LRR sequences.

A phylogenetic tree was constructed to investigate the relationship between the sequences described in this study and a selection of previously characterized soybean NBS-LRR sequences (Fig. 3). Included in this analysis were both sequences associated with the K644/B212 cluster of *R* genes (Jeong et al. 2001; Peñuela et al. 2002) and those located at chromosomal locations unlinked to *Rpg1-b* (Kanazin et al. 1996; Yu et al. 1996). In addition, we included several similar NBS-LRR sequences from common bean (*Phaseolus vulgaris*), some of which are linked to genes conferring resistance to *Colletotrichum lindemuthianum* (anthracnose) (Creusot et al. 1999; Ferrier-Cana et al. 2003).

With the exception of the H7 sequence, all of the soybean NBS-LRR sequences shown to be associated with the K644/B212 cluster (Fig. 3) form a well-supported group, suggesting a common ancestor. Within this group several subgroups with high bootstrap support are apparent. The four NBS-LRR sequences (RGA-39A, RGA-52A, RGA-84A, and NBSD-H8) present within the R45 and php2265 BAC contigs are all found within different, well-supported subgroups. It should be noted, however, that the NBS-LRR clones represented in this tree were isolated from several different cultivars and very similar sequences from different studies may be allelic. Significantly, the *P. vulgaris* sequences grouped tightly with the K644/B212 cluster of soybean genes, suggesting that this cluster was present prior to the divergence of the *Glycine* and *Phaseolus* genera. Furthermore, each of the four *P. vulgaris* sequences grouped with a separate soybean subgroup. Although the bootstrap support for these groupings was relatively weak, this pattern suggests that the subgroups may also predate formation of the *Glycine* genus.

Distribution of NBS-LRR families across the R45 and php2265 BAC contigs.

To determine the distribution of these NBS-LRR subgroups across our BAC contigs, we hybridized blots containing a contiguous set of BACs from the R45 and php2265 contigs, with probes derived from the NBS regions of the RGA-39A, RGA-52A, RGA-84A, and NBSD-H8 genes. These probes detected between two and four strongly hybridizing bands from the selected BACs (Fig. 4A). Under the stringency conditions used, the probes hybridized predominantly to distinct sets of bands, although some weak cross-hybridization was apparent (Fig.

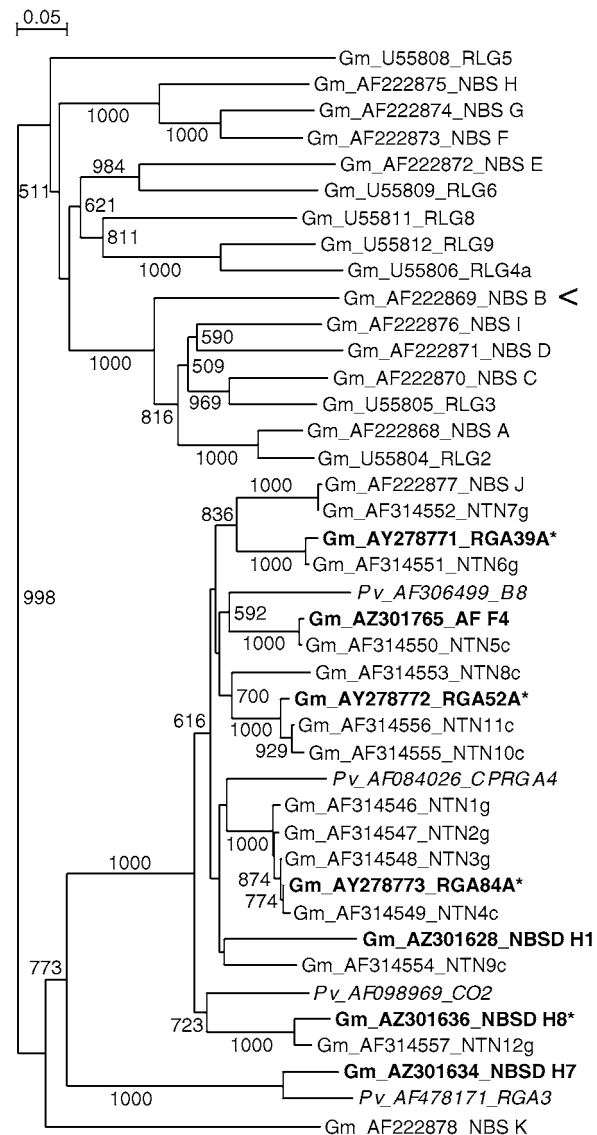


Fig. 3. Phylogenetic analysis of soybean nucleotide binding site-leucine rich repeat (NBS-LRR) sequences, including those associated with the *Rpg1-b* locus. The 258 nucleotides immediately following the P-loop were aligned from the indicated NBS-LRR clones. Sequences used in this study are shown in bold type. Sequences known to be represented within the BAC contigs described in this study are marked with asterisks. Coiled coil-NBS-LRR sequences associated with the K644/B212 cluster of *R* genes are identified with a vertical bracket on the right (excluding sequences from *Phaseolus vulgaris*). The toll/interleukin receptor-NBS-LRR sequence mapped to the K644/B212 cluster is marked with an arrow. Each sequence is identified with a prefix as either soybean (Gm) or common bean (Pv), followed by its GenBank accession number and alternative name. Alignments and trees were generated using Clustal X (Thompson et al. 1997). Maximum bootstrap values are 1,000. Bootstrap scores of less than 500 are not shown.

4A). NBSD-H1, NBSD-H7, and AF-F4 failed to hybridize strongly to any of the BACs (data not shown), indicating that only a subset of the NBS-LRR families associated with the K644/B212 cluster are found in the immediate vicinity of *Rpg1-b*.

In all, eleven putative NBS-LRR genes were identified, spread across the two BAC contigs (Fig. 4B). Five of these genes, belonging to either the RGA-52 or RGA-84 subgroups, were present on the *Rpg1-b* cosegregating BAC 52D1. There was a tendency for the NBS-LRR subgroups to be restricted in their physical distribution across the BAC contigs, indicative of tandem duplication events. This was especially striking for the RGA-52 family, with all four members detected being present within BAC 52D1 (insert size approximately 114 kb) (Fig. 4). However, two members of the RGA-84 subgroup are located on BAC 84D9, which is not contiguous with 52D1.

The RGA-52 and RGA-84 NBS-LRR families both include members that cosegregate with *Rpg1-b*. To determine the size of these NBS-LRR families in the soybean genome we hybridized the RGA-52A and RGA-84A probes to filters of *Hind*III digested genomic DNA from the cultivars Merit (*Rpg1-b*) and Flambeau (*rpg1-b*). Both probes detected small families of sequences with a high degree of polymorphism between the two

cultivars (Fig. 5). The four bands detected by the RGA-52A probe in cultivar Merit genomic DNA (Fig. 5) are similar in size to those detected in *Hind*III-digested BAC 52D1 (Fig. 4A), suggesting that the whole family may be present within BAC 52D1 (approximately 114 kb).

DISCUSSION

The resources available for high-resolution genetic and physical mapping in soybean have dramatically improved over the last few years with the generation of high-density PCR-based simple sequence length polymorphism (Cregan et al. 1999) and amplification fragment length polymorphism (Keim et al. 1997) genetic maps, the construction of several high-quality BAC libraries (Danesh et al. 1998; Marek and Shoemaker 1997; Meksem et al. 2000; Salimath and Bhattacharyya 1999; Tomkins et al. 1999), the genome-wide mapping of BACs (Fredrick Marek et al. 2001), and the generation of a large number of expressed sequence tags (Shoemaker et al. 2002). However, to our knowledge, there are no published reports of the map-based cloning of soybean disease resistance genes, and substantial hurdles remain. Soybean is an ancient tetraploid and has a complex genome in which many homologous regions are distributed throughout the

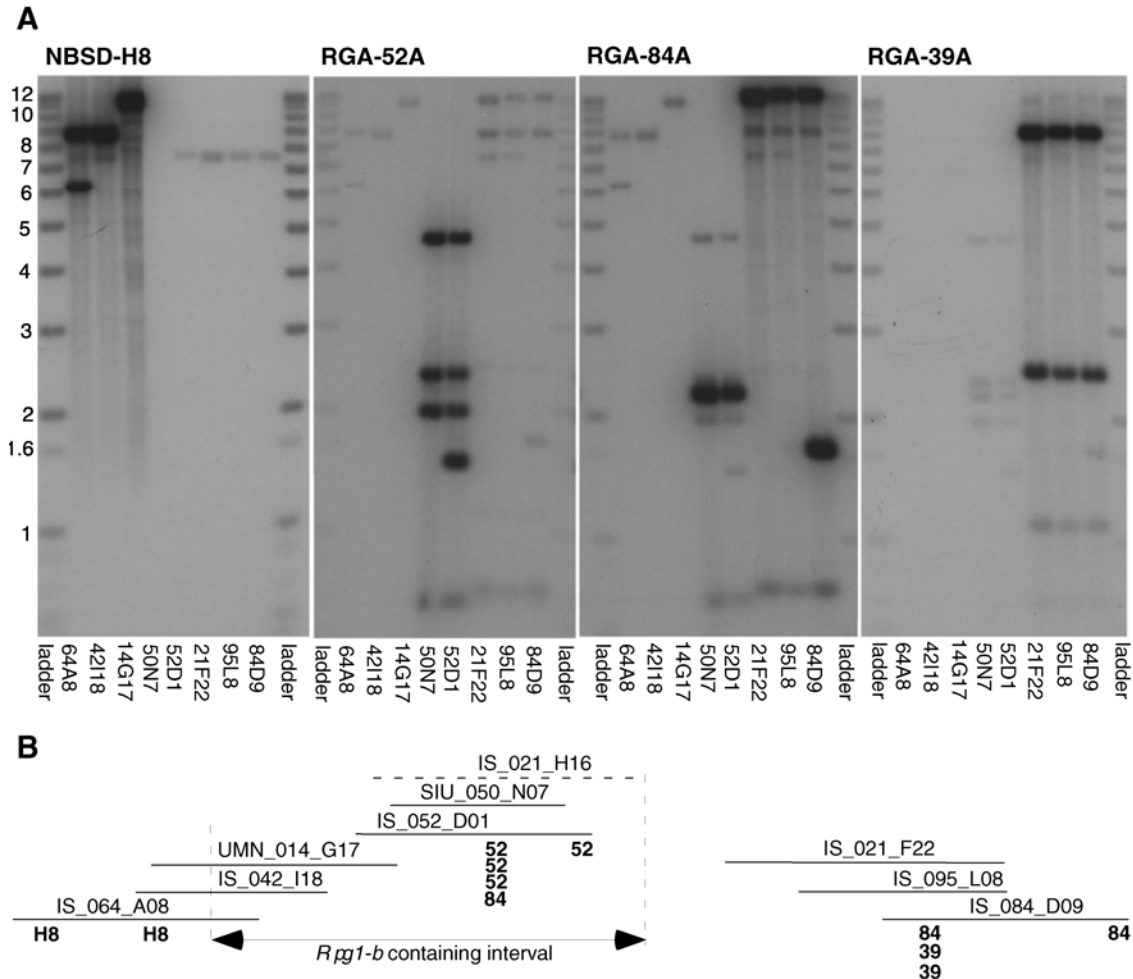


Fig. 4. Multiple families of nucleotide binding site-leucine rich repeat (NBS-LRR) sequences are associated with the *Rpg1-b* locus. **A**, Hybridization of NBS-LRR clones to the BAC contigs. Probes derived from the NBS region of the NBS-LRR clones RGA-39A, RGA-52A, RGA-84A, and NBSD-H8 were hybridized to filters of *Hind*III-digested BACs under high stringency conditions. BAC identities are indicated below each lane. The ladder band sizes are indicated in kb. **B**, Summary of the distribution of NBS-LRR gene families across the BAC contigs. BAC clones are indicated by horizontal lines. Numbers in bold indicate the presence of specific NBS-LRR sequences within the BAC. Different numbers are used to distinguish between NBS-LRR sequences detected by the different probes (e.g., sequences marked with a "52" were detected using the RGA-52A probe). BAC IS_021_H16 is marked with a dashed line to indicate that this BAC is not present on the blots shown in A. We have been unable to detect any additional NBS-LRR sequences within BAC IS_021_H16 (data not shown).

genome (Shoemaker et al. 1996). It has been reported that only 7.5% of RFLP markers detect single copy sequences (Shoemaker et al. 1996). The scarcity of single copy sequences limits the use of hybridization-based markers for high-resolution genetic and physical mapping in soybean. In this study, PCR-based markers were extensively used both for recombinant screening and for the screening of BAC libraries, due to their improved locus specificity.

The region immediately surrounding *Rpg1-b* is dramatically under-represented in the four soybean BAC libraries used in this study. Indeed, the region immediately proximal of the marker EP-21F22f is not represented in any of the libraries. This under-representation is probably explained by the instability and apparent toxicity to *E. coli* shown by BACs closely associated with the *Rpg1* locus.

The *Rpg1-b* gene is a member of a cluster of tightly linked resistance specificities located between RFLP markers K644 and B212 in molecular linkage group F. The cluster includes *R* genes and QTLs effective against bacterial, oomycete, fungal, viral, nematode, and insect pathogens (Ashfield et al. 1998; Diers et al. 1992; Gore et al. 2002; Rector et al. 1999; Roane et al. 1983; Tamulonis et al. 1997a and b). Two of these *R* genes, *Rsv1* and *Rpv1*, confer resistance to strains of soybean mosaic virus and peanut mottle virus, respectively. A high-resolution genetic map of the interval containing the *Rsv1* and *Rpv1* genes has recently been published (Gore et al. 2002). The use of common markers with the *Rpg1-b* genetic map presented here permits the partial integration of the two maps. It is clear from this analysis that all three *R* genes are very tightly linked, none mapping more than 0.6 cM from SSLP marker EP-64A8r (referred to as 64-A8C in the *Rsv1/Rpv1* map). The genetic data also indicates that *Rsv1* may reside within one of the BAC contigs described here, suggesting that our physical map may facilitate the characterization of other *R* genes within the cluster.

We show that the *Rpg1-b* locus is associated with at least four families of related CC-NBS-LRR sequences, with at least eleven members present within the BAC contigs. Five of these sequences cosegregate with *Rpg1-b*. Jeong and associates (2001) and Peñuela and associates (2002) have previously described several families of CC-NBS-LRR genes that map to the K644/B212 interval, and Gore and associates (2002) have placed a number of these on their high-resolution *Rsv1/Rpv1* map. The NBS-LRR sequences described in this study share strong sequence similarity with several of these genes, and the phylogenetic analysis described here demonstrates that most of the CC-NBS-LRR sequences associated with the K644/B212 *R* gene cluster fall within a single well-supported group. It appears that the *Rpg1-b* gene is part of a complex locus of many related CC-NBS-LRR sequences spread over several hundred kilobases and organized into small families of more closely related genes. Our observations are consistent with *Rpg1-b*, *Rsv1*, and *Rpv1* (and quite possibly other *R* genes in the cluster) being encoded by different members of these closely related CC-NBS-LRR-type *R* genes. There is a precedent for such an arrangement. Two highly similar *R* genes found within a 115-kb interval on potato chromosome 12 confer resistance to two unrelated pathogens, potato cyst nematode and potato virus X (van der Vossen et al. 2000). However, the majority of the clusters well characterized at the molecular level encode specificities restricted to different races of the same pathogen (Meyers et al. 1998; Parniske et al. 1997; Wei et al. 1999). The K644/B212 *R* gene cluster provides an excellent opportunity to characterize a cluster of closely related *R* genes conferring resistance to a wide array of unrelated pathogens.

The physical distribution of these NBS-LRR genes has been partially characterized by Peñuela and associates (2002), who identified 12 NBS-LRR genes located on six BAC contigs,

spread over an approximately 15-cM region of soybean MLG-F. The BAC clone 14G17 described in this study, which contains one member of the NBSD-H8 subgroup of NBS-LRR genes, was also identified by Peñuela and associates (2002) and was assigned to the J19 contig, allowing integration of these two physical maps. Significantly, our analysis identified three subgroups (RGA-52A, RGA-84A, and RGA-39A) comprising nine genes that were not identified by Peñuela and associates, indicating that there are likely many more NBS-LRR genes in the K644/B212 cluster yet to be uncovered.

The above observations are consistent with previous studies that have shown that disease resistance genes are often encoded by members of tightly-linked gene families (Hulbert et al. 2001). The organization of several complex *R* gene loci has been determined, and it is clear that the number of paralogs, their relatedness, and the physical distance separating them, varies greatly from locus to locus. Although sequences of highest similarity tend to be clustered together on the same BAC contig, this is not always the case, as evidenced by the RGA-84 family, copies of which are present in both the R45 and php2265 contigs (Fig. 4). A similar lack of correlation between physical position and level of sequence identity has also been observed in the *Dm3* cluster of lettuce (Meyers et al. 1998). Such distributions are indicative of relatively recent rearrangements within this *R* gene cluster and likely facilitate future rearrangements as recombination between highly similar copies should lead to local inversions and deletions. The extreme polymorphism between soybean cultivars Merit and Flambeau displayed by the RGA-52A subfamily (Fig. 5) is also indicative of a rapid rate of chromosomal rearrangements. It would be quite informative to compare the physical maps of this *R* gene cluster in these cultivars to assess the degree and type of rearrangements.

The presence of several distinct NBS-LRR subfamilies in close physical proximity suggests that gene conversion and re-

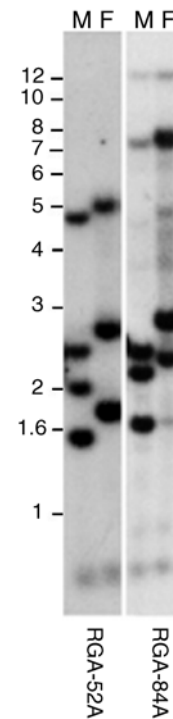


Fig. 5. RGA-52A and RGA-84A are members of small, highly polymorphic, gene families in the soybean genome. Shown are filters of *Hind*III-digested genomic DNA hybridized with probes derived from the nucleotide binding site (NBS) region of the indicated NBS-leucine rich repeat clones. M = cv. Merit (*Rpg1-b*), F = cv. Flambeau (*rpg1-b*). The positions of ladder bands are indicated in kb.

combination events between members of different subfamilies are relatively rare. This observation contrasts to that reported for the *Cf9* family of transmembrane LRR resistance genes in tomato, which appear to have undergone multiple sequence exchanges between tandemly repeated paralogs (Parniske and Jones 1999). The rate of recombination between *R* gene paralogs is likely influenced by both the physical separation between paralogs and the degree of nucleotide divergence. We are currently sequencing the entirety of BAC 52D1 to more accurately assess these factors.

The sequence of BAC 52D1 will also facilitate the definitive identification of *Rpg1-b*. To assist with this task, we have isolated an ethyl methanesulfonate-generated *rpg1-b* mutant allele (T. Ashfield and R. Innes, unpublished data) and are developing a transient transformation-based complementation assay similar to that described by Mindrinos and associates (1994).

The Arabidopsis *RPM1* gene has a related specificity to the soybean *Rpg1-b* gene, as both confer resistance to *Pseudomonas syringae* strains that express the avirulence gene *avrB* (Innes et al. 1993). It is not clear whether the *avrB* specificity has evolved more than once or whether a single ancestral gene with this specificity has been retained throughout the evolution of both Arabidopsis and soybean. The *RPM1* gene has been cloned and shown to be of the CC-NBS-LRR class (Grant et al. 1995) and is a simple locus consisting of a single NBS-LRR gene (Grant et al. 1995). In contrast, in this study, we demonstrate that the soybean *Rpg1-b* gene is part of a complex locus consisting of numerous CC-NBS-LRR-type genes. If *RPM1* and *Rpg1-b* evolved from the same ancestral gene, their loci have since taken very different evolutionary paths. Whatever the evolutionary relationship between *Rpg1-b* and *RPM1*, the cloning of *Rpg1-b* will permit the first direct comparison of two *R* genes with the same specificity from distantly related plant species.

MATERIALS AND METHODS

Plant lines and growth conditions.

The Flambeau (*rpg1-b*) × Merit (*Rpg1-b*) (Ashfield et al. 1995), Flyer (*Rpg1-b*) × Hartwig (*rpg1-b*) (Prabhu et al. 1999), and Essex (*rpg1-b*) × Forrest (*Rpg1-b*) (Hnetkovsky et al. 1996) RIL populations have been described previously. The Flambeau × Merit, Flyer × Hartwig, and Essex × Forrest populations were inbred to the F₈, F₅, and F₅, respectively. The soybean growth conditions used have been described elsewhere (Ashfield et al. 1995).

Disease assays.

The presence or absence of *Rpg1-b* was determined by hypersensitive response hand-inoculation tests, as described by Ashfield and associates (1995).

BAC libraries and screening.

The construction of the four BAC libraries used in this study has been described previously. The cv. Williams 82 library consists of approximately 40,000 clones with an average insert size of 150 kb and was constructed using genomic DNA partially digested with *HindIII* (Marek and Shoemaker 1997). The two cv. Forrest libraries contain approximately 38,000 clones each, with an average insert size of 125 kb and were constructed from *HindIII* or *BamHI* partial digests (Meksem et al. 2000). The cv. Faribault library contains approximately 31,000 clones with an average insert size of 120 kb and was constructed using an *EcoRI* partial digest (Danesh et al. 1998).

The Williams 82 library was screened by PCR using a three dimensional arrangement of BAC DNA pools (Marek and

Shoemaker 1997). The Forrest libraries were screened by PCR, essentially as described by Meksem and associates (2000). The Faribault library was screened using a combination of filter hybridizations and PCR. Primer pairs were first used to amplify the corresponding locus from genomic Faribault DNA. The amplified DNA was then purified, labeled with ³²P, and used to probe library filters (Danesh et al. 1998). Positive BAC clones were selected and rescreened by PCR, using the original primers.

BACs were isolated from the Williams 82 (Gm_ISb001), Faribault (Gm_UMNb001), and Forrest *HindIII* (Gm_SIUh001) libraries and are referred to in the figures with the prefixes IS-, UMN-, and SIU-, respectively. Elsewhere abbreviated BAC names are used that do not include the prefixes. BAC insert sizes were estimated from *HindIII* or *NorI* digests.

Isolating BAC DNA.

BAC DNA was isolated using Qiagen-tip 100 columns (Qiagen, Valencia, CA, U.S.A.), using procedures modified from those suggested by the manufacturer. Typically, the standard protocol was followed, except that the culture volume was increased to 100 ml and the volumes of P1, P2, and P3 buffers increased to 10 ml each. The elution buffer QF was prewarmed to 65°C, and the elution of the DNA from the column conducted in an oven at 65°C. The BAC DNA was resuspended in 50 µl of Tris-EDTA (TE) buffer. Larger yields could sometimes be achieved by using a culture volume of 500 ml and P1, P2, and P3 volumes of 40 ml each. In this instance, the cleared lysate obtained after P3 addition was precipitated with 0.7 volumes of room temperature isopropanol and the DNA was pelleted and then resuspended in 1.5 ml of TE. Before loading on the Qiagen-tip 100, the resuspended DNA was combined with 13.5 ml of QBT buffer to give a final volume of 15 ml. The purified BAC DNA was finally resuspended in 80 µl of TE.

Isolating plant genomic DNA.

Plant genomic DNA for Southern blot analyses was isolated using the DNeasy Plant Kit (Qiagen), using the protocol supplied by the manufacturer. DNA for PCR was isolated essentially as described by Bisgrove and associates (1994). A small leaf sample (typically an immature leaf approximately 1 cm in length) was transferred to a 1.5-ml microfuge tube, was frozen in liquid nitrogen, was crushed using a microfuge pestle, and 300 µl of cetyltrimethylammonium bromide (CTAB) buffer (2% CTAB, 100 mM Tris-HCl, pH 8, 20 mM EDTA, 1.4 M NaCl, 1% polyvinyl polypyrrolidone MW40000) was added. The sample was then vortexed and placed at 65°C for at least 10 minutes before being extracted with an equal volume of chloroform/isoamyl alcohol (24:1). The DNA was precipitated from the aqueous layer by the addition of 3 volumes of ethanol and was pelleted. The pellet was washed with 1 ml of 70% ethanol, and finally, was resuspended in 200 µl of TE containing 50 µg of RNaseA per ml. One to two microliters was used per 20 µl of PCR reaction.

Digests, preparation of filters, and hybridizations.

Southern blotting and filter hybridizations were as described by Ashfield and associates (1998). Hybridizations were at 60°C and washes at 55°C. The NBS probes correspond to the region between the P-loop and NBSD domains.

PCR amplifications.

The contents of PCR reactions were as described by Ashfield and associates (1998). Large-scale PCR experiments were set up in 96-well microtiter plates and were run in a PTC-100 (MJ Research, Waltham, MA, U.S.A.) thermal cycler. After a 2-min denaturation at 94°C, the reactions were cycled 32 times, using the profile 94°C for 45 s, 52 to 56°C for 1 min,

and 72°C for 1 min. The specific annealing temperature was determined by the primer pair. Amplified products were typically separated through 3% MetaPhor agarose (Cambrex Bio Science, Rockland, ME, U.S.A.).

Linkage analysis.

Map distances in the RILs were calculated using the Haldane and Waddington equation, as described by Ashfield and associates (1995).

Isolating BAC end-sequences.

BAC end sequences were obtained by IPCR, by plasmid rescue, or by direct BAC end sequencing. IPCR and plasmid rescue were conducted essentially as described by Woo and associates (1994). See below for details on BAC end sequencing.

Sequencing and sequence analysis.

Sequencing reactions were prepared using the ABI Big Dye Terminator cycle sequencing kit (Applied Biosystems, Foster City, CA, U.S.A.), as described by the manufacturer with some modifications. PCR products and small plasmids were sequenced in 10- μ l reactions containing 2 μ l of Terminator Ready Reaction Mix, 2 μ l of 5-mM MgCl₂, 1.5 pmol of primer, and between 10 to 100 ng of template. BAC templates were sequenced in 40- μ l reactions containing 16 μ l of Terminator Ready Reaction Mix, 7 pmol of primer, and at least 0.5 μ g of template. All reactions were run on an ABI3700, and traces were examined using the Sequencher software (GeneCodes, Ann Arbor, MI, U.S.A.).

Generating molecular markers.

Primers were developed from genomic DNA sequences (RFLP markers R45 and php2265 and BAC end sequences), and the corresponding loci amplified from the three pairs of mapping parents. In some instances, length polymorphisms (LPs) were apparent between products without digestion. If LPs were not apparent, the amplified products were purified and sequenced. Sequences were compared for single-nucleotide polymorphisms (SNPs). SNPs that created or removed restriction sites were converted to cleaved amplified polymorphic sequence markers by digestion of the PCR products with the appropriate restriction enzyme. When necessary, SNPs were scored by amplifying and sequencing the polymorphic sequence from each relevant mapping line. A table describing all primers and polymorphisms present in each cultivar is available from R. Innes upon request.

ACKNOWLEDGMENTS

The authors acknowledge V. Njiti, Department of Agriculture, Alcorn State University, MS, U.S.A. for supplying seed of the E \times F and H \times F populations. This work was supported by United States Department of Agriculture (USDA)-NRICGP grants 95-37300-1589 and 99-35300-7693 to R. W. Innes, National Science Foundation (NSF) grant 9872635, and Illinois Soy Program Operating Board grants 95-222-24-3 and 98-222-25-2 to D. A. Lightfoot and K. Meksem, and USDA-CSREE grant 98-3599-6168 and NSF grant DBI-9872664 to N. D. Young.

LITERATURE CITED

Arahana, V. S., Graef, G. L., Specht, J. E., and Steadman, J. R. 2001. Identification of QTLs for *Sclerotinia* stem rot resistance in soybean using SSR markers. *Crop Sci.* 41:180-188.

Ashfield, T., Danzer, J. R., Held, D., Clayton, K., Keim, P., Saghai Maroof, M. A., Webb, P. M., and Innes, R. W. 1998. *Rpg1*, a soybean gene effective against races of bacterial blight, maps to a cluster of previously identified disease resistance genes. *Theor. Appl. Genet.* 96:1013-1021.

Ashfield, T., Keen, N. T., Buzzell, R. I., and Innes, R. W. 1995. Soybean resistance genes specific for different *Pseudomonas syringae* avirulence genes are allelic, or closely linked, at the *RPG1* locus. *Genetics* 141:1597-1604.

Bisgrove, S. R., Simonich, M. T., Smith, N. M., Sattler, A., and Innes, R. W. 1994. A disease resistance gene in Arabidopsis with specificity for two different pathogen avirulence genes. *Plant Cell* 6:927-933.

Cregan, P. B., Jarvik, T., Bush, A. L., Shoemaker, R. C., Lark, K. G., Kahler, A. L., Kaya, N., VanToai, T. T., Lohnes, D. G., Chung, J., and Specht, J. E. 1999. An integrated genetic linkage map of the soybean genome. *Crop Sci.* 39:1464-1490.

Creusot, F., Macadre, C., Ferrier Cana, E., Riou, C., Geffroy, V., Seignac, M., Dron, M., and Langin, T. 1999. Cloning and molecular characterization of three members of the NBS-LRR subfamily located in the vicinity of the Co-2 locus for anthracnose resistance in *Phaseolus vulgaris*. *Genome* 42:254-264.

Danesh, D., Peñuela, S., Mudge, J., Denny, R. L., Nordstrom, H., Martinez, J. P., and Young, N. D. 1998. A bacterial artificial chromosome library for soybean and identification of clones near a major cyst nematode resistance gene. *Theor. Appl. Genet.* 96:196-202.

Dangl, J. L., and Jones, J. D. 2001. Plant pathogens and integrated defence responses to infection. *Nature* 411:826-833.

Diers, B. W., Mansur, L., Imsande, J., and Shoemaker, R. C. 1992. Mapping phytophthora resistance loci in soybean with restriction fragment length polymorphism markers. *Crop Sci.* 32:377-383.

Ferrier-Cana, E., Geffroy, V., Macadre, C., Creusot, F., Imbert-Bollere, P., Seignac, M., and Langin, T. 2003. Characterization of expressed NBS-LRR resistance gene candidates from common bean. *Theor. Appl. Genet.* 106:251-261.

Flor, H. H. 1955. Host-parasite interactions in flax rust-its genetics and other implications. *Phytopathology* 45:680-685.

Fredrick Marek, L., Mudge, J., Darnielle, L., Grant, D., Hanson, N., Paz, M., Huihuang, Y., Denny, R., Larson, K., Foster-Hartnett, D., Cooper, A., Danesh, D., Larsen, D., Schmidt, T., Staggs, R., Crow, J. A., Retzel, E., Young, N. D., and Shoemaker, R. C. 2001. Soybean genomic survey: BAC-end sequences near RFLP and SSR markers. *Genome* 44:572-581.

Gore, M. A., Hayes, A. J., Jeong, S. C., Yue, Y. G., Buss, G. R., and Saghai Maroof, M. A. 2002. Mapping tightly linked genes controlling potyvirus infection at the *Rsv1* and *Rpv1* region in soybean. *Genome* 45:592-599.

Grant, M. R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R. W., and Dangl, J. L. 1995. Structure of the Arabidopsis *RPM1* gene enabling dual specificity disease resistance. *Science* 269:843-846.

Hayes, A. J., and Saghai Maroof, M. A. 2000. Targeted resistance gene mapping in soybean using modified AFLPs. *Theor. Appl. Genet.* 100:1279-1283.

Hayes, A. J., Yue, Y. G., and Saghai Maroof, M. A. 2000. Expression of two soybean resistance gene candidates shows divergence of paralogous single-copy genes. *Theor. Appl. Genet.* 101:789-795.

Hnetkovsky, N., Chang, S. J. C., Doubler, T. W., Gibson, P. T., and Lightfoot, D. A. 1996. Genetic mapping of loci underlying field resistance to soybean sudden death syndrome (SDS). *Crop Sci.* 36:393-400.

Hulbert, S. H., and Bennetzen, J. L. 1991. Recombination at the *Rp1* locus of maize. *Mol. Gen. Genet.* 226:377-382.

Hulbert, S. H., Webb, C. A., Smith, S. M., and Sun, Q. 2001. Resistance gene complexes: Evolution and utilization. *Annu. Rev. Phytopathol.* 39:285-312.

Innes, R. W., Bisgrove, S. R., Smith, N. M., Bent, A. F., Staskawicz, B. J., and Liu, Y. C. 1993. Identification of a disease resistance locus in Arabidopsis that is functionally homologous to the *RPG1* locus of soybean. *Plant J.* 4:813-820.

Islam, M. R., Shepherd, K. W., and Mayo, G. M. E. 1989. Recombination among genes at the *L* group in flax conferring resistance to rust. *Theor. Appl. Genet.* 77:540-546.

Jeong, S. C., Hayes, A. J., Biyashev, R. M., and Saghai Maroof, M. A. 2001. Diversity and evolution of a non-TIR-NBS sequence family that clusters to a chromosomal "hotspot" for disease resistance genes in soybean. *Theor. Appl. Genet.* 103:406-414.

Kanazin, V., Fredrick Marek, L., and Shoemaker, R. C. 1996. Resistance gene analogs are conserved and clustered in soybean. *Proc. Natl. Acad. Sci. U.S.A.* 93:11746-11750.

Keim, P., Schupp, J., Travis, S., Clayton, K., Zhu, T., Shi, L., Ferreira, A., and Webb, D. 1997. A high-density soybean genetic map based on AFLP markers. *Crop Sci.* 37:537-543.

Marek, L. F., and Shoemaker, R. C. 1997. BAC contig development by fingerprint analysis in soybean. *Genome* 40:420-427.

Meksem, K., Zobrist, K., Ruben, E., Hyten, D., Quanzhou, T., Zhang, H.-B., and Lightfoot, D. A. 2000. Two large-insert soybean genomic li-

- barries constructed in a binary vector: Applications in chromosome walking and genomic wide physical mapping. *Theor. Appl. Genet.* 101:747-755.
- Meyers, B. C., Chin, D. B., Shen, K. A., Sivaramakrishnan, S., Lavelle, D. O., Zhang, Z., and Michelmore, R. W. 1998. The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. *Plant Cell* 10:1817-1832.
- Meyers, B. C., Dickerman, A. W., Michelmore, R. W., Sivaramakrishnan, S., Sobral, B. W., and Young, N. D. 1999. Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J.* 20:317-332.
- Michelmore, R. W., and Meyers, B. C. 1998. Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* 8:1113-1130.
- Mindrinos, M., Katagiri, F., Yu, G. L., and Ausubel, F. M. 1994. The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* 78:1089-1099.
- Pan, Q., Wendel, J., and Fluhr, R. 2000. Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. *J. Mol. Evol.* 50:203-213.
- Parniske, M., Hammond-Kosack, K. E., Golstein, C., Thomas, C. M., Jones, D. A., Harrison, K., Wulff, B. B. H., and Jones, J. D. G. 1997. Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. *Cell* 91:821-832.
- Parniske, M., and Jones, J. D. 1999. Recombination between diverged clusters of the tomato Cf-9 plant disease resistance gene family. *Proc. Natl. Acad. Sci. U.S.A.* 96:5850-5855.
- Peñuela, S., Danesh, D., and Young, N. D. 2002. Targeted isolation, sequence analysis, and physical mapping of non-TIR NBS-LRR genes in soybean. *Theor. Appl. Genet.* 104:261-272.
- Prabhu, R. R., Njiti, V. N., Bell-Johnson, B., Johnson, J. E., Schmidt, M. E., Klein, J. H., and Lightfoot, D. A. 1999. Selecting soybean cultivars for dual resistance to soybean cyst nematode and sudden death syndrome using two DNA markers. *Crop Sci.* 39:982-987.
- Pryor, T., and Ellis, J. 1993. The genetic complexity of fungal resistance genes in plants. *Adv. Plant Pathol.* 10:281-305.
- Rector, B. G., All, J. N., Parrott, W. A., and Boerma, H. R. 1999. Quantitative trait loci for antixenosis resistance to corn earworm in soybean. *Crop Sci.* 39:531-538.
- Roane, C. W., Tolin, S. A., and Buss, G. R. 1983. Inheritance of reaction to two viruses in the soybean cross York x Lee 68. *J. Hered.* 74:289-291.
- Salimath, S. S., and Bhattacharyya, M. K. 1999. Generation of a soybean BAC library and identification of DNA sequences tightly linked to the *Rps-k* disease resistance gene. *Theor. Appl. Genet.* 98:712-720.
- Shoemaker, R. C., Polzin, K., Labatte, J., Specht, J., Brummer, E. C., Olson, T., Young, N., Concibido, V., Wilcox, J., Tamulonis, J. P., Kochert, G., and Boerma, H. R. 1996. Genome duplication in soybean (*Glycine subgenus soja*). *Genetics* 144:329-338.
- Shoemaker, R., Keim, P., Vodkin, L., Retzel, E., Clifton, S. W., Waterston, R., Smoller, D., Coryell, V., Khanna, A., Erpelding, J., Gai, X., Brendel, V., Raph-Schmidt, C., Shoop, E. G., Vielweber, C. J., Schmatz, M., Pape, D., Bowers, Y., Theising, B., Martin, J., Dante, M., Wylie, T., and Granger, C. 2002. A compilation of soybean ESTs: Generation and analysis. *Genome* 45:329-338.
- Stahl, E. A., Dwyer, G., Mauricio, R., Kreitman, M., and Bergelson, J. 1999. Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. *Nature* 400:667-671.
- Tamulonis, J., Luzzi, B., Hussey, R., Parrott, W., and Boerma, H. 1997a. DNA markers associated with resistance to Javanese root-knot nematode in soybean. *Crop Sci.* 37:783-788.
- Tamulonis, J. P., Luzzi, B. M., Hussey, R. S., Parrott, W. A., and Boerma, H. R. 1997b. DNA marker analysis of loci conferring resistance to peanut root-knot nematode in soybean. *Theor. Appl. Genet.* 95:664-670.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876-4882.
- Tomkins, J. P., Mahalingam, R., Smith, H., Goicoechea, J. L., Knap, H. T., and Wing, R. A. 1999. A bacterial artificial chromosome library for soybean PI 437654 and identification of clones associated with cyst nematode resistance. *Plant Mol. Biol.* 41:25-32.
- van der Vossen, E. A., van der Voort, J. N., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D. C., Bakker, J., Stiekema, W. J., and Klein-Lankhorst, R. M. 2000. Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: A virus and a nematode. *Plant J.* 23:567-576.
- Wei, F., Gobelmann-Werner, K., Morroll, S. M., Kurth, J., Mao, L., Wing, R., Leister, D., Schulze-Lefert, P., and Wise, R. P. 1999. The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley. *Genetics* 153:1929-1948.
- Witsenboer, H., Kesseli, R. V., Fortin, M. G., Stanghellini, M., and Michelmore, R. W. 1995. Sources and genetic structure of a cluster of genes for resistance to three pathogens in lettuce. *Theor. Appl. Genet.* 91:178-188.
- Yu, Y. G., Buss, G. R., and Saghai Maroof, M. A. 1996. Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. *Proc. Natl. Acad. Sci. U.S.A.* 93:11751-11756.
- Yu, Y. G., Saghai Maroof, M. A., Buss, G. R., Maughan, P. J., and Tolin, S. A. 1994. RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. *Phytopathology* 84:60-64.