Recombination Within a Nucleotide-Binding-Site/Leucine-Rich-Repeat Gene Cluster Produces New Variants Conditioning Resistance to Soybean Mosaic Virus in Soybeans

A. J. Hayes,*[,1] S. C. Jeong,* M. A. Gore,*[,2] Y. G. Yu,*[,3] G. R. Buss,* S. A. Tolin[†] and M. A. Saghai Maroof*[,4]

**Department of Crop and Soil Environmental Sciences and [†]Department of Plant Pathology, Physiology and Weed Science, Virginia Tech, Blacksburg, Virginia 24061*

Manuscript received April 24, 2003
Accepted for publication October 5, 2003

ABSTRACT

The soybean *Rsv*1 gene for resistance to soybean mosaic virus (SMV; Potyvirus) has previously been described as a single-locus multi-allelic gene mapping to molecular linkage group (MLG) F. Various *Rsv*1 alleles condition different responses to the seven (G1–G7) described strains of SMV, including extreme resistance, localized and systemic necrosis, and mosaic symptoms. We describe the cloning of a cluster of NBS-LRR resistance gene candidates from MLG F of the virus-resistant soybean line PI96983 and demonstrate that multiple genes within this cluster interact to condition unique responses to SMV strains. In addition to cloning 3gG2, a strong candidate for the major *Rsv*1 resistance gene from PI96983, we describe various unique resistant and necrotic reactions coincident with the presence or absence of other members of this gene cluster. Responses of recombinant lines from a high-resolution mapping population of PI96983 (resistant) × Lee 68 (susceptible) demonstrate that more than one gene in this region of the PI96983 chromosome conditions resistance and/or necrosis to SMV. In addition, the soybean cultivars Marshall and Ogden, which carry other previously described *Rsv*1 alleles, are shown to possess the 3gG2 gene in a NBS-LRR gene cluster background distinct from PI96983. These observations suggest that two or more related non-TIR-NBS-LRR gene products are likely involved in the allelic response of several *Rsv*1-containing lines to SMV.

THE traditional model of a single dominant gene conferring resistance to a specific pathogen species or subspecies has been described in numerous plant/pathogen systems (for a review see Ellis *et al.* 2000). Specific genes from important crop plants and model plant systems that confer resistance to viruses, fungi, bacteria, and nematodes have been cloned and characterized. Analysis of these genes at the sequence level has shown that many different resistance genes from divergent plant species encode proteins containing very similar structural components.

A major class of cloned resistance genes encode an N-terminal nucleotide-binding-site (NBS) domain and a C-terminal leucine-rich-repeat (LRR) region. More than 20 plant disease resistance genes from this class have been cloned, including important crop resistance genes such as the virus resistance gene *N* from tobacco (Whitham *et al.* 1994), the virus resistance gene *Rx* from potato (Bendahmane *et al.* 1999), the maize rust resistance gene *Rp*1-D (Collins *et al.* 1999), and the rice blast resistance gene *Pi*b (Wang *et al.* 1999). In addition, genome and expressed sequence tag sequencing projects, along with directed PCR-based approaches, have demonstrated that NBS-LRR resistance gene candidate (RGC) sequences are prevalent in plant genomes (Kanazin *et al.* 1996; Leister *et al.* 1996; Yu *et al.* 1996; Meyers *et al.* 1999; Pan *et al.* 2000).

In soybean, the *Rsv*1 resistance gene is tightly linked to a cluster of NBS-LRR RGCs on the molecular linkage group (MLG) F (Yu *et al.* 1996). *Rsv*1 confers resistance to soybean mosaic virus, an important potyvirus pathogen of soybean. Several alleles at this locus have been reported, including *Rsv*1 from PI96983, *Rsv*1-y from York, *Rsv*1-t from Ogden, and *Rsv*1-m from Marshall (Cho and Goodman 1979). These alleles are distinguished by their differential response to the seven strains of soybean mosaic virus (SMV; G1–G7) as characterized by Cho and Goodman (1979). The resistance response observed with *Rsv*1 is similar to the *Rx*-mediated extreme resistance to potato virus X (Bendahmane

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AY518517–AY518519.

[1]*Present address:* Philip Morris, P.O. Box 26603, Richmond, VA 23261.

[2]*Present address:* Department of Analytical Biochemistry, DuPont Agriculture and Nutrition, 7300 NW 62nd Ave., P.O. Box 1004, Johnston, IA 50131-1004.

[3]*Present address:* Eli Lilly and Co., BioResearch Technologies and Proteins, Drop Code 0444, Lilly Corporate Center, Indianapolis, IN 46285.

[4]*Corresponding author:* Department of Crop and Soil Environmental Sciences, 328 Smyth Hall, Virginia Tech, Blacksburg, VA 24061-0404. E-mail: smaroof@vt.edu

et al. 1997) and *Ry*-mediated extreme resistance to potato virus Y in potato (Hamalainen *et al.* 1997): localized cell death is not observed and virus cannot be detected in the inoculated leaf. In addition, several *Rsv*1/SMV strain interactions result in systemic necrosis and plant death 2–3 weeks after inoculation (Chen *et al.* 1991). In this study, we focus on the *Rsv*1 allele from PI96983 that confers extreme resistance to strains G1–G6 and systemic lethal necrosis to strain G7.

Yu *et al.* (1994) mapped the *Rsv*1 allele of PI96983 to a disease resistance gene cluster on MLG F. Using PCR-based approaches, several partial- and full-length NBS-LRR RGCs have been localized to the MLG F near *Rsv*1 (Yu *et al.* 1996; Hayes and Saghai Maroof 2000; Jeong *et al.* 2001; Peñuela *et al.* 2002). Yu *et al.* (1996) identified two classes of NBS sequence (classes b and j) that map to this resistance gene cluster. The class b NBS, represented by the full-length clone L20a (Hayes *et al.* 2000), is a single-copy gene that belongs to the subgroup of NBS-LRR genes possessing a Drosophila Toll/human interleukin-1 receptor (TIR) homologous region (for a review of NBS-LRR classification, see Young 2000). The class j NBS RGCs, which belong to the non-TIR subclass, make up a large family of homologous sequences clustered at or near the *Rsv*1 locus (Jeong *et al.* 2001; Gore *et al.* 2002; Peñuela *et al.* 2002). Because many highly homologous, tightly linked class j genes and/or pseudogenes colocalize to MLG F, mapping individual gene members with respect to a disease resistance locus requires developing techniques to unambiguously detect single gene members and then map these genes to independent loci.

The goal of this research was to determine if any of the class j non-TIR-NBS-LRR candidate genes present on the MLG F of soybean represent the *Rsv*1 disease resistance locus. In this report, we describe the cloning and sequence characterization of a strong candidate for the major *Rsv*1 locus in PI96983 and provide evidence that multiple genes within the soybean non-TIR-NBS-LRR gene cluster on MLG F contribute to the plant's response to SMV.

MATERIALS AND METHODS

Plant genetic materials: The source of the *Rsv*1 gene (allele) in this study is the soybean line PI96983. DNA samples from PI96983 and the Williams isoline L81-4420, which possesses *Rsv*1 resistance derived from PI96983, were used for cloning and sequence analyses. Additionally, we conducted SMV reaction studies and DNA genotyping on the SMV-susceptible cultivars Williams and Lee 68 and the *Rsv*1-containing differential lines, Marshall, Ogden, York, Kwanggyo, and PI507389.

Genetic mapping: The segregating population of PI96983 (*Rsv*1) × Lee 68 (*rsv*1) was used for mapping and identification of recombinant lines. Numerous markers have been previously mapped to the disease resistance gene cluster on MLG F in this population of 1056 individuals (Gore *et al.* 2002). Briefly, all 1056 $F_{2:3}$ lines were genotyped for three PCR-based molecular markers. Two microsatellite markers, hsp176 and SATT120, flank the Rsv1 locus by 2.9 and 3.8 cM, respectively. The third marker, 64A8c, maps 0.5 cM away from the *Rsv*1 locus. These three markers were used to identify $F_{2:3}$ lines carrying a recombination in the region of interest. Disease genotypes for F_2 individuals having undergone a recombination in the 6.7-cM region were determined on the basis of a greenhouse screening of 15–20 $F_{2:3}$ individuals from each F_2 line with SMV strain G1. These selected lines were further genotyped with multiple high-resolution markers, including class G gene probes. Fourteen rare $F_{2:3}$ recombinant families that had undergone a crossover in the 1.3-cM gene cluster near *Rsv*1 were advanced in the field by an additional selfed generation. The resultant F_4 seed of individual plants that were homozygous for the recombinant chromosome region near *Rsv*1 were collected for additional genotyping and SMV reaction studies to confirm gene (marker) order and to further implicate candidate SMV resistance genes.

SMV disease reaction: Seeds of the soybean parental lines PI96983 and Lee 68 and F_4 homozygous recombinants were planted in the greenhouse in 15-cm pots containing Metro Mix commercial potting media. Approximately 10 days after planting, fully expanded unifoliolate leaves were inoculated with SMV according to the method of Hunst and Tolin (1982). Individual recombinant F_4 lines were evaluated with SMV strains G1–G3 and G5–G7 (G4 was not available for testing). Recorded reactions included symptomless (resistant), mosaic (susceptible), and various necrotic reactions, including localized necrotic spots and veinal necrosis on the inoculated unifoliolate leaf and systemic necrosis of the noninoculated trifoliolate leaves. Differential cultivars of known SMV-strain response were included in the disease reaction studies to verify the identities of the SMV strains and effectiveness of inoculation.

Screening of genomic clones: Using class j sequence-specific probes, we screened genomic libraries of the soybean lines Williams 82 (SMV susceptible) and L81-4420 (a Williams isoline that contains chromosomal contribution for the resistance gene cluster on MLG F from PI96983). The Williams 82 genomic library was obtained from Stratagene (La Jolla, CA) and was constructed in the λFIXII vector. Hybridization was performed as detailed in Hayes *et al.* (2000). This library was used for the initial identification of class j genomic sequences. Large λ-subclones from the Williams 82 library were digested with restriction enzymes *Dde*I and *Hin*fI. Restriction fragments were subcloned and then used as probes for Southern analysis on a group of soybean lines to identify low-copy probes.

Isolated probes gG, eG, and gGsp (identified in this study) were used for subsequent screening of the custom-made L81-4420 genomic library. This was done to identify specific class j gene sequences in a line containing the *Rsv*1 gene. This genomic library was constructed from young leaf tissue by CLONTECH (Palo Alto, CA) in the EMBL SP6/T7 vector and was screened according to the manufacturer's protocols. Positive λ-clones were digested with *Sst*I and fragments were subcloned into the pBluescript KS(−) vector for further analysis.

Southern hybridization: Class j gene-specific probes gG, gGsp, and eG (all three probes identified in this study) were hybridized to Southern-blotted genomic DNA from parents and homozygous recombinant progeny of the cross PI96983 × Lee 68 and from other soybean *Rsv*1 differential cultivars. DNA was extracted from young leaf tissue according to Yu *et al.* (1994) and Gore *et al.* (2002). Southern blotting was performed as described in Hayes *et al.* (2000).

PCR amplification of 3gG2: The PI96983 *Rsv*1 gene candidate, 3gG2 (identified in this study), was PCR amplified from DNA of cultivars Marshall and Ogden using two pairs of nested primers. A PCR product with an expected size of ~4 kb was

amplified using the primers 3gG2-5′A (5′-TTACTCATAAAAT TAAAACAAACAAGTG-3′) and 3gG2-3′A (5′-ATCATATTTC AAAAGCTACAAAGATACC-3′). For PCR amplification, 150 ng of genomic DNA was used as template in a 50-μl reaction containing 1× reaction buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3); 4 mM $MgCl_2$; 2 μM of each primer; 800 μM each of dNTPs; 2.5 units of *Taq* polymerase. Thirty-four cycles of a PCR reaction were run with denaturation at 94° for 30 sec, primer annealing at 48° for 30 sec, and primer extension at 72° for 150 sec. This PCR reaction resulted in many products being amplified in addition to the expected size product. Therefore, the diluted PCR reaction was used as template for a second round of PCR. A 3.3-kb internal fragment was amplified using the primers 3gG2-5′B (5′-ATCCACTCAACTCCCTTTCC TACC-3′) and 3gG2-3′B (5′-ACGCGTTTAATGTGAGCAATCT-3′). For PCR amplification, 2 μl of 100-fold diluted primary PCR reaction was used as template in a 50-μl reaction containing 1× reaction buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3); 2.5 mM $MgCl_2$; 1 μM of each primer; 400 μM each of dNTPs; 2.5 units of *Taq* polymerase. Thirty-four cycles of a PCR reaction were run with denaturation at 94° for 30 sec, primer annealing at 55° for 30 sec, and primer extension at 72° for 150 sec.

DNA sequence analysis: DNA sequencing of small plasmid and PCR template was performed using the Dye Terminator cycle sequencing kit and an ABI 377 automated DNA sequencer according to the manufacturer's protocols (Perkin-Elmer, Foster City, CA). Large plasmid subclones were sequenced by transposon-mediated shotgun sequencing using the EZ::TN <TET-1> transposon insertion kit (Epicentre, Madison, WI). Sequence editing, contig construction, multiple-sequence alignment, and primer design were conducted using Lasergene software from DNAstar (Madison, WI).

RESULTS

Within the class j RGC gene family on MLG F, we identified a subclass of six non-TIR-NBS-LRR genes (or pseudogenes) in PI96983 (hereafter, referred to as subclass G). All of these RGCs possess highly conserved sequence in the 3′-untranslated region. The G subclass was identified by evaluating the genomic clone G27. G27 was isolated from a Williams 82 (which lacks SMV resistance) genomic library probed with a class j non-TIR-NBS-LRR cDNA probe containing the 5′ portion of a resistance gene candidate sequence. The G27 genomic clone, which contains a full-length class j open reading frame (ORF) sequence, hybridizes to numerous fragments on soybean genomic Southern blots. This hybridization pattern, consisting of several polymorphic bands, multiple monomorphic bands, and comigrating fragments, is typical of all the clones we studied containing the conserved class j ORF sequence. However, several small DNA fragments of the G27 genomic clone, derived from outside the class j gene-coding region (see MATERIALS AND METHODS), revealed low complexity, highly polymorphic restriction fragment length polymorphism patterns when hybridized to the same genomic Southern blots. The 1.5-kb probe gG, located just 3′ of the translated region, detects six subclass G sequence family members (represented by the genomic clones 1eG30, 5gG3, 3gG2, 1eG15, 6gG9, and 1gG4) in PI96983 (Figure 1, A and B, and Table 1). By exploiting a 500-bp indel, we developed a second probe, gGsp, that detects only three (3gG2, 1eG15, and 6gG9) of the six subclass G members in PI96983 (Figure 1, A and B). A third probe of 850 bp, eG, is only partially characterized at the sequence level in the subclass G genes and was used primarily for mapping purposes. This sequence is located ~3 kb downstream of the end of translation (Figure 1A) and detects at least five of the six PI96983 gene family members (hybridization not shown). All of the six subclass G gene members detected in PI96983 map to a tightly clustered region of the soybean MLG F near *Rsv*1 and no other loci are detected with these probes (Figure 1B; for a more detailed map, see GORE *et al.* 2002).

A λ-genomic library constructed from the *Rsv*1 isoline L81-4420 was screened with the probes gG, gGsp, and eG. A total of 35 genomic clones were isolated. Not surprisingly, most clones were detected with more than one of the three probes. Candidate λ-clones were digested with *Sst*I and the digested fragments of 16 selected clones were subcloned into the pBluescript vector. On the basis of restriction digestion pattern and/or by end-sequencing individual *Sst*I subclones, we were able to classify all 35 λ-genomic clones into one of the six subclasses. This analysis is based on restriction fragment and sequence identity around a conserved *Sst*I site in the 3′-untranslated region of the subclass G genes (shown in Figure 1A). This grouping corresponds directly to the number of fragments detected on a *Hin*dIII-digested Southern blot (Figure 1B, gG probe). Genomic subclones 3gG2 (8 kb, AY518517), 5gG3 (8 kb, AY518518), and 6gG9 (10 kb, AY518519), representing three of the six candidate genes, were completely sequenced, and this sequence information was used to confirm the identity of those restriction fragments in *Hin*dIII Southern blots. Genomic clones possessing the remaining three candidate genes were only partially sequenced to confirm unique identities. Identities of the restriction fragments for these three clones on *Hin*dIII Southern blots of PI96983 genomic DNA were verified by digesting the full-length λ-clones with *Hin*dIII and probing with the gG clone.

***Rsv*1 gene candidate 3gG2:** Map locations of subclass G genes were determined in the PI96983 × Lee 68 population of 1056 individuals (GORE *et al.* 2002). Map positions were verified first by selecting individual homozygous recombinants for each line possessing a recombination in the subclass G region and then by confirming both the genotypes and disease responses of these lines (discussed in detail below). On the basis of restriction digestion data and sequence-predicted restriction fragment size, we confirmed the *Hin*dIII restriction fragment corresponding to the genomic clone 3gG2. To the exclusion of all other known class j genes, the gene candidate represented by the 3gG2 genomic clone cosegregates with the PI96983 *Rsv*1 locus condi-

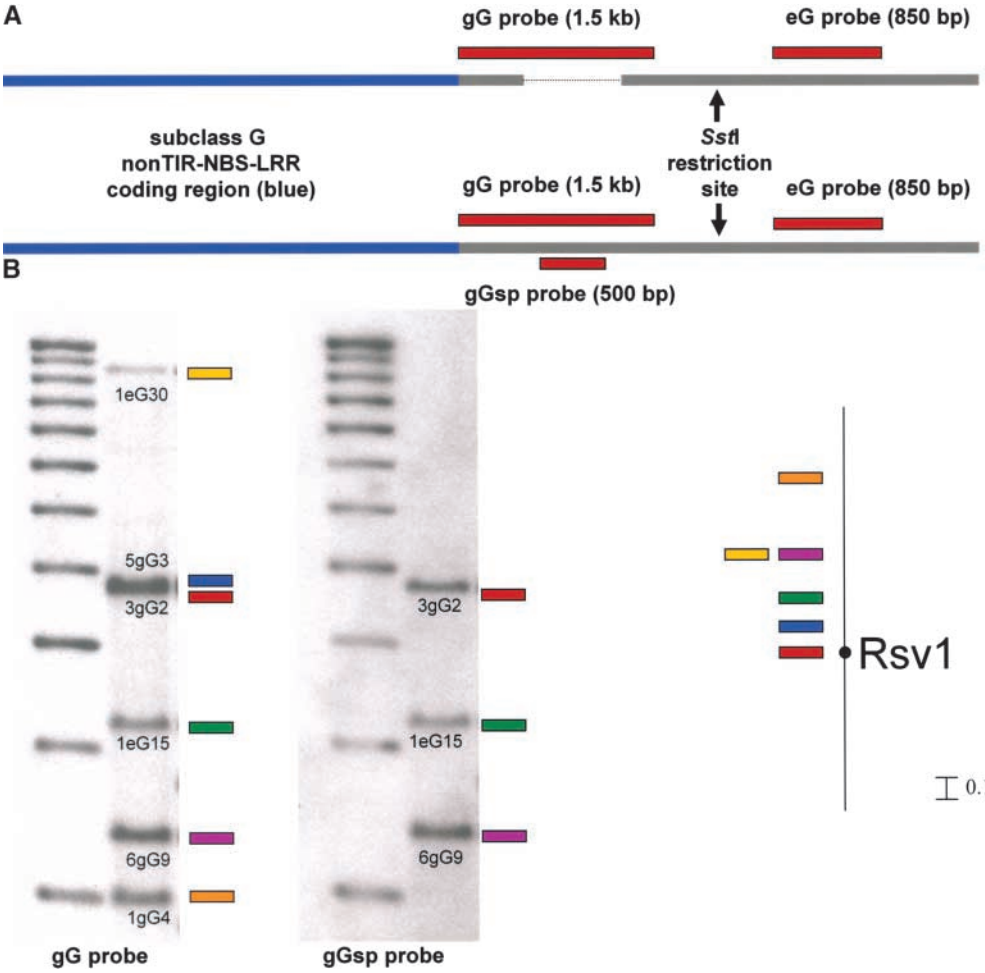

FIGURE 1.—Description of the subclass G family, including (A) sequence features for an ~8-kb region common to members of the subclass G gene family from soybean line PI96983 (the blue shaded area is the coding region and the light areas show the 3′ noncoding region). Bars represent clones without (top) and with (bottom) the 500-bp indel (not drawn to scale). The three probes discussed in the text are derived from genomic regions 3′ of the NBS-LRR class j gene candidate coding region. The gGsp probe is a 500-bp indel present in three of the six subclass G members from PI96983. The conserved *Sst*I site was used as an anchor point for deriving sequence information from numerous clones to classify the subclass G genes into six divisions. (B) Autoradiograph of PI96983 Southern blotted genomic DNA digested with *Hin*dIII (right lane of each film) probed with gG (left film) and gGsp (right film). The left lane of each film is a 1-kb ladder. The bottom band of the ladder corresponds to 2 kb and the band sizes increase in 1-kb increments. Colored boxes to the right of each band and names below each band correspond to a single λ-genomic clone assigned to each detected size fragment on the basis of sequence data of *Hin*dIII restriction sites and band sizes of *Hin*dIII-digested λ-clones probed with gG and gGsp. The map to the right depicts the corresponding genetic location of each fragment with respect to the major *Rsv*1 resistance locus as determined in the high-resolution mapping population PI96983 (*Rsv*1) × Lee 68 (*rsv*1).

tioning extreme resistance to SMV strain G1 and systemic necrosis to SMV strain G7.

The λ-genomic clone containing this gene candidate is ~20 kb and the completely sequenced 3gG2 subclone, 7979 bp, incorporates 2465 bp of 5′-untranslated sequence and 2125 bp of 3′-untranslated sequence. The 3gG2 ORF sequence encodes a 3390-bp gene, whose deduced protein product is highly similar to previously cloned non-TIR-NBS-LRR disease resistance genes. A BLASTx of the National Center for Biotechnology Information nonredundant protein database showed that the deduced amino acid sequence most closely resembles a family of non-TIR-NBS-LRR sequences found near the *Co*-2 locus of the common bean (*Phaseolus vulgaris*; CREUSOT *et al.* 1999). 3gG2 is also highly similar to the *I*2-C gene of tomato, which conditions resistance to a Fusarium wilt disease (ORI *et al.* 1997). The BLASTx E-values for these hits were 0.0 and e^{-173}, respectively. A phylogenetic tree of sequences related to 3gG2 is shown in Figure 2. RT-PCR products from total RNA of PI96983 leaf tissue was isolated and a nearly full-length cDNA of the 3gG2 gene was completely sequenced, thus demonstrating that the 3gG2 gene is expressed in leaf tissue (data not shown).

TABLE 1

Six class G genomic clones

Class G genomic clone	*Hin*dIII fragment size (kb)	Approximate distance from *Rsv1* (cM)
1eG30	10.5	0.5
5gG3	4.81	0.1
3gG2	4.76	0.0
1eG15	3.2	0.2
6gG9	2.5	0.5
1gG4	2.0	0.9

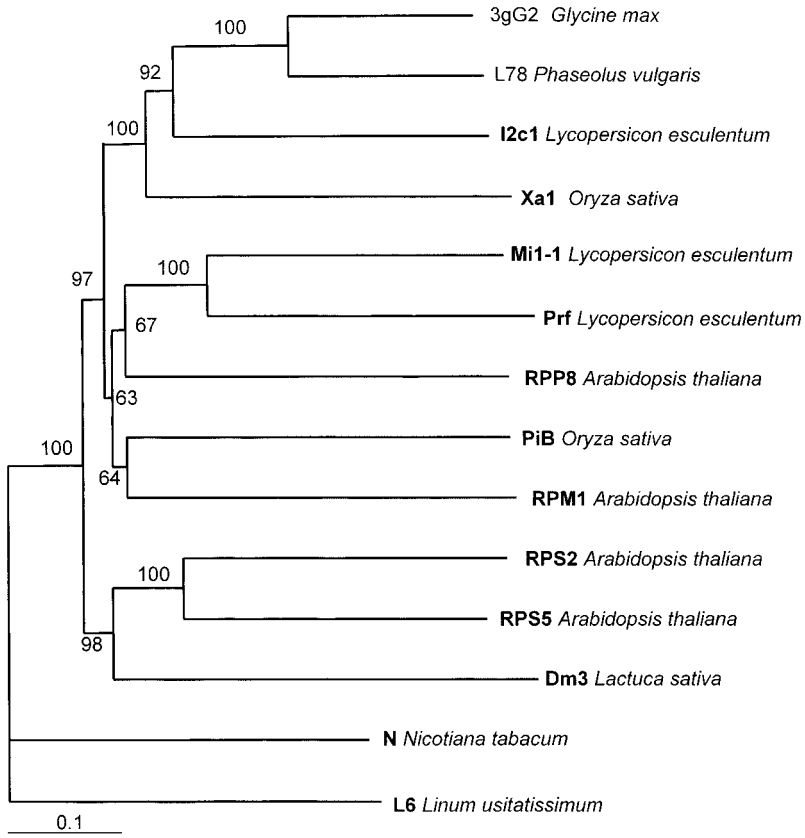

FIGURE 2.—Phylogenetic relationship of the 3gG2 gene to known disease resistance genes (boldface type) as well as resistance gene candidates from previous studies. The phylogenetic tree was constructed by the neighbor-joining method (SAITOU and NEI 1987) as implemented in PAUP* (SWOFFORD 2001). The tree is based on the alignment of amino acid sequences containing the nucleotide-binding-site motif and its surrounding conserved regions, which correspond to the amino acid sequence positions 204–525 from the N terminus of the 3gG2 gene product. The branch lengths are proportional to the average substitutions per site as indicated by the scale. Values of 100 bootstrap replicates are indicated on branches. Genes are labeled by their recognized gene name followed by their species of origin.

Disease screening of homozygous recombinants: In developing the high-resolution mapping population for placing these subclass G gene candidates in the F linkage group, we identified 14 rare recombinants that had undergone crossovers within the subclass G gene family. Initially, we evaluated the disease reaction in segregating $F_{2:3}$ lines. However, in response to inoculation with SMV strain G1, combinations of resistant, necrotic, and susceptible plants were observed in single segregating lines. These reactions were observed only in the lines possessing a heterozygous recombination in the subclass G gene region and are postulated to be a gene dosage effect of one or several genes from the *Rsv*1 region of PI96983.

To study the recombinants in a more homogenous background, $F_{2:3}$ seed were planted in the field. Individual plants were genotyped with tightly flanking simple sequence repeat markers to select a single plant from each line that was homozygous for the recombination in the subclass G gene region. The F_4 progeny of these homozygous recombinant plants were tested for their SMV disease reaction to various SMV strains (representatives shown in Table 2). These F_4 homozygous recombinants are expected to be still segregating at 12.5% of the loci outside of the gene cluster on linkage group F. To confirm that unlinked epistatic interactions were not contributing to the disease response, we replicated disease reactions as many as four times, with 4–10 plants tested in each reaction. We also conducted tests at two separate greenhouses with two different groups of investigators. These SMV reaction studies on numerous segregating individuals demonstrate that resistance, necrosis, and susceptible reactions are conditioned exclusively by a locus or tightly linked loci on MLG F (GORE *et al.* 2002).

Among the homozygous recombinants studied, lines that carry the 3gG2 gene show extreme resistance to SMV-G1. This includes the line 800-46 that carries only the 3gG2 gene from PI96983 along with five subclass G genes from Lee 68. Inoculation with SMV-G7 in these same lines results in systemic necrosis similar to that observed in PI96983. Several homozygous recombinants that do not possess the 3gG2 gene show initial resistance to strain G1, followed by late necrosis in some individuals. This includes the line 943-9 that carries five of the six subclass G genes from PI96983 but does not carry 3gG2 (Figure 3) or any subclass G genes from Lee 68. Recombinant lines not carrying 3gG2 are consistently susceptible to G7. It was not possible to specifically assign this observed reaction to a particular mapped locus, but it is clear that at least one additional gene tightly linked to 3gG2 conditions a resistance response to SMV strain G1. We are currently conducting complementation tests to confirm the role of 3gG2 in SMV resistance and also to test other subclass G members for response to SMV.

TABLE 2

Class G gene composition and reaction of soybean differential cultivars and recombinant F_4 lines (from the cross PI96983 × Lee68) to various SMV strains

Recombinant F_4 lines or cultivars	Class G genes from PI96983	General reaction to inoculation with SMV strains					
		SMV-G1	SMV-G2	SMV-G3	SMV-G5	SMV-G6	SMV-G7
800-46	3gG2	R	R	LLN	LLN	R	SN
943-9	1eG30 5gG3 1eG15 6gG9 1gG4	R Some late Necrosis	R	R	S	S	S
613-10	1eG30 3gG2 5gG3 1eG15 6gG9	R	R	LLN	LLN	R	SN
1044-98	1eG30 3gG2 5gG3 1eG15 6gG9	R	R	LLN	LLN	R	SN
PI96983	1eG30 3gG2 5gG3 1eG15 6gG9 1gG4	R	R	R	R	R	SN
Lee 68	?	S	S	S	S	S	S
Marshall	3gG2 ?	R	R	LLN	LLN	R	SN
Ogden	3gG2 ?	R	R	LLN	LLN	R	SN

R, resistant; S, susceptible; SN, systemic necrosis; LLN, local lesion necrosis; "?," presence of additional class G genes that were not characterized as part of this study.

Detailed disease phenotyping of six rare homozygous recombinant lines with SMV strains G1–G3 and G5–G7 show that various strain-specific virus reactions completely unique from the parental lines are observed. In particular, several recombinant lines, including 613-10 and 1044-98, show localized necrotic lesions followed by systemic necrosis when inoculated with strains G3 and G5 (Figure 4A). PI96983 shows complete resistance to these strains and never gives localized lesions. Another recombinant, 800-46, possesses the 3gG2 gene but none of the other subclass G genes derived from PI96983. On the basis of Southern blotting evidence, this line contains all five subclass G genes from the susceptible parent, Lee 68, in addition to the 3gG2 gene from PI96983 (Figure 3, gG probe). This line shows complete resistance to strain G1, while strains G3 and G5 produce local lesions followed by systemic necrosis. Strain G7 inoculation of 800-46 leads to rapid systemic necrosis, which in four replicated experiments was shown to occur in about half the time as the same reaction in PI96983 (Figure 4B). All of these lines demonstrate the occurrence of new *Rsv*1 specificities to SMV strains coincident with recombination of the subclass G gene complement.

The 3gG2 gene candidate in Marshall and Ogden: Several rare homozygous recombinants from the studies described above show reactions to SMV closely resembling those of the soybean differential cultivars Marshall and Ogden (*e.g.*, G3 and G5 reactions in Figure 4A). Marshall and Ogden possess two previously characterized *Rsv*1 alleles, *Rsv*1-m and *Rsv*1-t, respectively (Chen *et al.* 1991). Southern hybridization with the gG probe indicates that Marshall and Ogden possess the 3gG2 gene found in PI96983. This is based on an identical-sized band that is present in all three lines (Figure 3). To confirm this postulation, we specifically amplified the 3gG2 gene from Marshall and Ogden by PCR using two pairs of nested primers (see materials and methods). We also attempted to amplify the 3gG2 gene from the susceptible cultivars, Lee 68 and Williams, but no PCR product was amplified, offering further support that this specific member of the G subclass is not present in these lines. Sequence analysis of the PCR products from Marshall and Ogden demonstrates that both possess the identical putative resistance gene, 3gG2, which is associated with resistance in PI96983 (sequence data not shown).

Interestingly, the Southern hybridization pattern with the probes gG and gGsp indicates that, with the exception of 3gG2, the subclass G gene family complements for Marshall and Ogden, which are identical to one another, differ considerably from PI96983 (Figure 3). Furthermore, these hybridization patterns more closely resemble several homozygous recombinants derived from this study, including 613-10 and 1044-98 (Figure 3). The phenotypic reaction of these recombinant lines

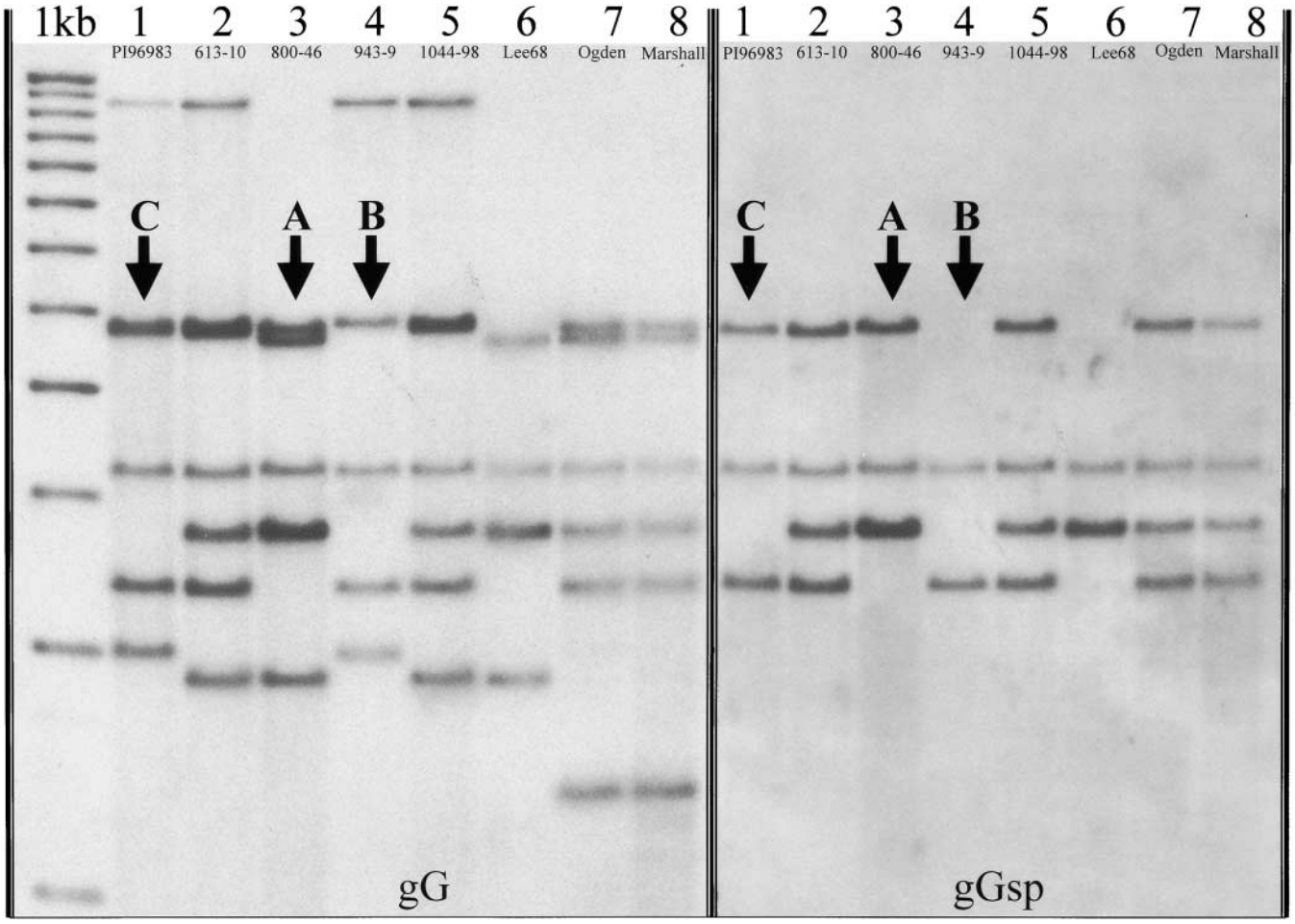

FIGURE 3.—Autoradiograph of Southern blotted soybean genomic DNA digested with *Hin*dIII and probed with gG (left). The blot was stripped and then probed with gGsp (right). Lanes 1 and 6 are the parents PI96983 (*Rsv*1) and Lee68 (*rsv*1), respectively, of the high-resolution mapping population from which the F_4 rare homozygous recombinants 613-10 (lane 2), 800-46 (lane 3), 943-9 (lane 4), and 1044-98 (lane 5) are derived. Lanes 7 and 8 contain DNA from SMV-resistant soybean lines Ogden (*Rsv*1-t) and Marshall (*Rsv*1-m). The 1-kb ladder consists of a 1-kb bottom band with subsequent fragments increasing in size by 1-kb increments. Arrow A shows the *Hin*dIII fragment representing 3gG2 (4742 bp) that is present in 800-46 (not 943-9) and is detected by the full-length gG probe and the indel probe gGsp (the lower band in the doublet detected by the gG probe is derived from the Lee 68 parent, lane 6). Arrow B shows the *Hin*dIII fragment representing 5gG3 (4806 bp) present in 943-9 (not 800-46) that is detected solely by the full-length gG probe. Arrow C shows that in the PI96983 parent, the gG probe detects both fragments (double intensity) and the gGsp probe detects only the 3gG2 fragment (single intensity). On the basis of segregation of homozygous F_4 progeny, Lee 68 is purported to carry five subclass G genes, three of which possess the 500-bp indel. Hybridization with the gG and gGsp probes shows a double intensity band at 2.8 kb that is also present in the progeny line 800-46. 613-10 and 1044-98 have only a single intensity band at 2.8 kb.

to various SMV strains closely resembles the responses observed in Marshall and Ogden (*e.g.*, Figure 4A). The homozygous recombinants 613-10 and 1044-98 have subclass G genes derived from both the susceptible Lee 68 parent and the resistant PI96983 parent, thus leading to an overall genotype and phenotype that differs markedly from either parent. It should be emphasized that the vast majority of recombinations in the PI96983 × Lee68 mapping population occur outside the subclass G gene family and that the predicted SMV disease reaction in these lines segregates as a single genetic locus. Only through high-resolution mapping efforts did we identify rare segregants in which the subclass G genes of PI96983 and Lee 68 had been shuffled.

DISCUSSION

The soybean MLG F contains 15 or more NBS-LRR gene candidate sequences (JEONG *et al.* 2001; PEÑUELA *et al.* 2002). All but one of these known NBS sequences belong to a highly homologous non-TIR-NBS-LRR gene family referred to as class j by YU *et al.* (1996). By exploiting conserved sequences in the 3′ noncoding region of several of these class j genes, we developed a

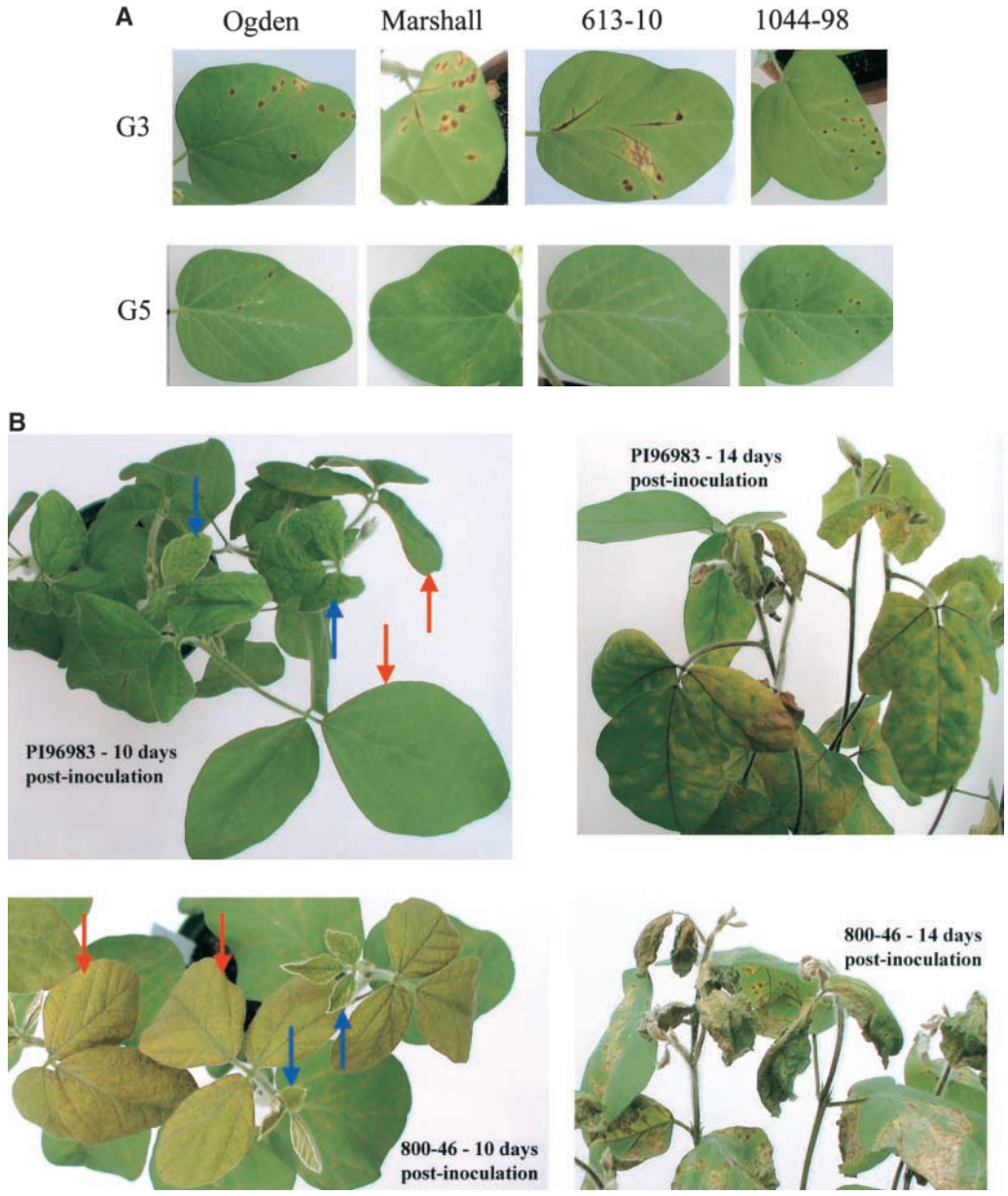

FIGURE 4.—Pictures of soybean reaction following SMV inoculation. (A) Cultivars Ogden (*Rsv*1-t) and Marshall (*Rsv*1-m) and recombinant lines 613-10 and 1044-98 (derived from this study) show large localized necrotic lesions and veinal necrosis in response to inoculation with SMV strain G3. Inoculation with strain G5 produces similar but less pronounced lesions. All photographs are from plants of a single replication at 18 days postinoculation. (B) Soybean line PI96983 and recombinant line 800-46 (derived from this study) show systemic necrosis in response to inoculation with strain G7, 10 and 14 days postinoculation. At 10 days, 800-46 shows severe bronzing of the first trifoliolate (systemic) leaf (red arrows) as well as stunting and necrosis of the younger trifoliolates (blue arrows). By 14 days, trifoliolate leaves are completely necrotic. At 10 days, PI96983 shows a mild response in the first trifoliolate leaf (red arrows), followed by more pronounced necrosis in younger trifoliolate leaves (blue arrows). By 14 days, all trifoliolate leaves show moderate to severe necrosis.

set of probes that enabled us to specifically "fingerprint" an important group (subclass G) of six class j genes (Figure 1). The subclass G genes are tightly linked to the SMV resistance gene, *Rsv*1. By matching mapping data of bands of particular sizes from the gG and gGsp probes with the specific corresponding restriction sites of genomic clones, we were able to confirm the chromosomal location for five of the six subclass G gene members in PI96983. The sixth band (1eG15) is monomorphic in the PI96983 × Lee 68 population; however, we have putatively deduced the location of the sixth subclass G member on the basis of mapping data of the eG probe (located 3 kb downstream of the gG sequence). Because of the complexities associated with cross detection in studying multigene families, discerning each of the individual subclass G genes was critical to identifying 3gG2, a strong candidate to be the major gene conferring resistance to SMV. Of the six members of the G subclass within the class j gene family from PI96983, 3gG2, 5gG3, and 6gG9 have been completely characterized at the sequence level. Representative clones of the three remaining subclass G members were only partially sequenced to verify identity. It was confirmed that both 3gG2 and 5gG3 are expressed in the leaf tissue of PI96983 by sequencing of RT-PCR products (data not shown). Conversely, 6gG9 was never detected in expression studies.

Several groups have isolated bacterial artificial chromosomes (BACs) from the region of soybean linkage group F near *Rsv*1 (PEÑUELA *et al.* 2002; R. INNES, personal communication). However, estimates of physical distance have been difficult to determine in this region because of genetic mapping complexities associated with cross-detection. On the basis of colocalization of

subclass G clones on Faribault BACs (kindly provided by N. Young) we conservatively estimate the physical distance in the *Rsv*1 genomic region to be between 225 and 450 kb/cM. The 20-kb 3gG2 genomic clone carries a single resistance gene candidate and our best evidence from exhaustive genomic-library surveys and high-resolution genetic mapping is that this is the only RGC cosegregating with the major *Rsv*1 locus. However, ongoing complementation studies will provide the necessary evidence to confirm the purported function of the 3gG2 gene.

Soybean mosaic virus resistance is conditioned by a complex gene family: Virus reaction studies of several homozygous recombinants that have undergone crossovers in the subclass G gene region demonstrate conclusively that SMV resistance in PI96983 is conditioned by a complex of tightly linked genes. While the major SMV reaction is likely controlled by the 3gG2 gene, it is apparent that one or more additional genes within the subclass G gene region of PI96983 also condition a response to SMV. Confirmation of this observation is evidenced by the fact that several homozygous recombinant lines demonstrate resistance and/or necrotic responses to SMV strains G1, G2, and G3 despite the absence of the 3gG2 gene. This reaction is clearly unique from the extreme resistance observed in lines carrying the 3gG2 gene because late necrosis is observed in some plants inoculated with strain G1 and strains G5–G7 produce susceptible mosaic symptoms. Further studies show that several homozygous recombinants that do not possess the full complement of subclass G genes, but do possess the 3gG2 gene, show localized necrosis to several strains, including G3 and G5. Since this type of reaction is not observed in the PI96983 parent or in any other progeny lines that do not have recombinations in the subclass G gene region, it can be concluded only that multiple genes within this region complement each other in responding to SMV inoculation. Cho and Goodman (1982) reported unique necrotic reactions in several plants from a group of advanced soybean breeding lines possessing SMV resistance derived from PI96983. Our recent findings suggest that these observations were likely the result of rare recombinations within the tightly linked subclass G family.

Sun *et al.* (2001) recently reported that variant resistance reactions to *Puccinia sorghi* conferred by the *Rp*1-D locus of maize result from haplotype changes as a result of intergenic recombination. We have demonstrated that previously reported allelism at the *Rsv*1 locus, at least among several soybean lines, is the product of similar variations in haplotype resulting from intragenic recombination. Ogden and Marshall carry the complete 3gG2 resistance gene found in PI96983. However, these lines respond to strains G3 and G5 in a markedly different fashion from that of PI96983. Several homozygous recombinants developed in this study carry 3gG2 along with other subclass G genes derived from the susceptible parent, Lee 68. These plants show SMV disease reactions closely resembling those of Ogden and Marshall. It has been previously reported that two unlinked genes in Arabidopsis, RXC3 from Ler and RXC4 from Col-0, condition interdependent resistance to *Xanthomonas campestris*, when combined in a single genotype (Buell and Somerville 1997). In soybean, we have demonstrated that "allelism" at the *Rsv*1 locus, which was previously believed to be controlled by a single dominant gene, is actually conditioned by a similar multi-genic interaction in which tightly linked members of a multigene family contribute to the host pathogen interaction.

Implications for candidate gene cloning: By isolating and characterizing several homozygous recombinant F_4 plants from the cross PI96983 × Lee 68, we demonstrated that 3gG2 alone cannot condition the specific resistance observed in PI96983 for all SMV strains. This observation is critical in predicting the outcome of future complementation studies. We do not expect susceptible lines transformed with 3gG2 alone to show extreme resistance to each of the six strains, G1–G6, as observed in PI96983. However, lines carrying 3gG2 consistently confer systemic necrosis to strain G7 and extreme resistance to strain G1; thus these reactions should be good indicators that the appropriate candidate gene has been cloned.

In studying the model Arabidopsis-*Pseudomonas syringae* host-pathogen system, Banerjee *et al.* (2001) recently observed that the Po-1 allele of the NBS-LRR resistance gene, *Rps*2, confers resistance in a Col-0 background despite the fact that this allele does not condition resistance in its native Po-1 genotype. They postulated that additional host factors contribute to the ability of the *Rps*2 gene product to effectively elicit a hypersensitive response. It is clear that transacting factors cannot explain the major phenomena (*i.e.*, extreme resistance or lethal necrosis) observed in the *Rsv*1/SMV system, given the observed uniform reaction in the heterogenous fixed recombinants across replications. Although we cannot rule out the presence of tightly linked host factors contributing to the overall SMV reaction, the level of genetic resolution and the phenotypic uniformity in SMV reactions observed among lines not carrying a crossover in the small chromosomal region around *Rsv*1 strongly suggest that multiple members of the subclass G gene family contribute to the various SMV resistance responses. It is interesting to note, however, that the late necrosis to G1, observed in several homozygous recombinant lines not carrying 3gG2, was not observed in all the plants tested from these lines. In this case, the presence of additional unlinked host factors or unrelated environmental factors could explain the variable response.

Evolutionary aspects: It has previously been demonstrated that soybean carries numerous duplicate loci on homeologous chromosomes, lending credence to its purported ancient tetraploid origin (Shoemaker *et al.*

1996). For example, the soybean resistance gene candidate, L20a, which also resides on MLG F near *Rsv*1, has a paralog on the soybean MLG E (Hayes *et al.* 2000). Similarly, class j gene-candidate sequences have been mapped to MLGs E and A2, in addition to the multiple loci detected on MLG F near *Rsv*1 (Jeong *et al.* 2001). By contrast, it is interesting to note that the subclass G genes all reside in a tight cluster on the F linkage group and that no other paralogs carrying this conserved 3′-untranslated sequence are detected in the genome. Thus, this particular subclass within the class j gene family likely arose since the divergence of tetraploid soybean from its diploid progenitor. Furthermore, in the PI96983 line, three gene members carry a 500-bp indel in the 3′ flanking region and three do not. This observation, combined with map positions for the six candidate genes, suggests that duplication followed by divergence (*i.e.*, insertion or deletion) and further duplication may explain the six-gene cluster observed in PI96983.

Southern blotting evidence and homozygous recombinant segregation data indicate that the susceptible cultivar from the mapping population in this study, Lee 68, carries only five subclass G gene members (see Figure 3). Sequence data from 35 genomic clones demonstrate that PI96983 carries six subclass G genes, and hybridization of Marshall and Ogden DNA with the gG probe strongly suggests that each carries six subclass G genes also. When DNA from 800-46 is probed with gG, all five bands from Lee 68 are detected in addition to the 3gG2 band from PI96983. Conversely, for 943-9 DNA, gG hybridizes to five bands derived from PI96983, but does not detect 3gG2 or any Lee 68 bands (Figure 3, gG probe, lanes 3 and 4). One explanation for this observation is that no homolog to the 3gG2 gene is present in Lee 68. Grant *et al.* (1995) observed an analogous phenomenon at the *Rpm*1 resistance locus in Arabidopsis in which no homolog exists in the susceptible line. They suggest that either the gene has arisen recently or reduction has occurred through genomic instability.

However, the observation that Lee 68 likely carries three loci where the 500-bp indel gGsp is present, identical to PI96983, which carries three loci including 3gG2 (Figure 3, lanes 1 and 6, gGsp probe), suggests an alternative explanation. A cross between haplotypes of similar structure has the potential for mispairing and unequal crossing over because of cognate features (Michelmore and Meyers 1998). Evidence from the *Rp*1-D gene cluster suggests reduction due to unequal crossing over (Collins *et al.* 1999; Sun *et al.* 2001). Recombinant 800-46 carries four loci detected by gGsp, and 943-9 carries only two loci detected by gGsp (Figure 3, lanes 3 and 4). This evidence suggests that unstable pairing may occur at this gene family location and that unequal crossing over could be a major contributing factor in the creation of new haplotypes. In the *Rsv*1/SMV pathosystem, we observe that new arrangements of gene family members interact in unique ways to condition novel responses to SMV infection. Thus, the purported allelism that has been reported in this host-pathogen system (Cho and Goodman 1979) may be more complex than was originally thought. Rather than resulting from variations of a single gene locus, the *Rsv*1 allelic series is due to rearrangements of several genes leading to novel soybean haplotype-SMV pathotype specificities.

Compounding evidence suggests that the resistance-gene-rich region on MLG F near *Rsv*1 is a highly variable region. Evidence from this subclass G gene study supports previous reports that rapid evolution of resistance gene clusters is an important source of variation (Richter and Ronald 2000). Our future studies will focus on characterizing subclass G genes from other resistant soybean genotypes and defining exactly which genes interact to condition variable responses to SMV.

We thank N. Young and S. Peñuela for BAC Southern blots of class j BAC clones and Moss Baldwin for technical assistance. This study was supported in part by the United States Department of Agriculture National Research Initiative Competitive Grants Program grant no. 96-35300-3648, the United Soybean Board, and the Virginia Soybean Board.

LITERATURE CITED

Banerjee, D., X. Zhang and A. Bent, 2001 The leucine-rich repeat domain can determine effective interaction between *RPS*2 and other host factors in Arabidopsis *RPS*2-mediated disease resistance. Genetics **158:** 439–450.

Bendahmane, A., K. Kanyuka and D. C. Baulcombe, 1997 High-resolution genetical and physical mapping of the *Rx* gene for extreme resistance to potato virus X in tetraploid potato. Theor. Appl. Genet. **95:** 153–162.

Bendahmane, A., K. Kanyuka and D. C. Baulcombe, 1999 The *Rx* gene from potato controls separate virus resistance and cell death responses. Plant Cell **11:** 781–791.

Buell, C. R., and S. C. Somerville, 1997 Use of *Arabidopsis* recombinant inbred lines reveals a monogenic and novel digenic resistance mechanism to *Xanthomonas campestris* pv. *campestris*. Plant J. **12:** 21–29.

Chen, P., G. R. Buss, C. W. Roane and S. A. Tolin, 1991 Allelism among genes for resistance to soybean mosaic virus in strain-differential soybean cultivars. Crop Sci. **31:** 305–309.

Cho, E. K., and R. M. Goodman, 1979 Strains of soybean mosaic virus: classification based on virulence in resistant soybean cultivars. Phytopathology **69:** 467–470.

Cho, E. K., and R. M. Goodman, 1982 Evaluation of resistance in soybeans to soybean mosaic virus strains. Crop Sci. **22:** 1133–1136.

Collins, N., J. Drake, M. Ayliffe, Q. Sun, J. Ellis *et al.*, 1999 Molecular characterization of the maize *Rp*1-D rust resistance haplotype and its mutants. Plant Cell **11:** 1365–1376.

Creusot, F., C. Macadré, E. Ferrier Cana, C. Riou, V. Geffroy *et al.*, 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.

Ellis, J., P. Dodds and T. Pryor, 2000 Structure, function, and evolution of plant disease resistance genes. Curr. Opin. Plant Biol. **3:** 278–284.

Gore, M. A., A. J. Hayes, S. C. Jeong, Y. G. Yue, G. R. Buss *et al.*, 2002 Mapping tightly linked genes controlling *potyvirus* infection at the *Rsv*1 and *Rpv*1 region in soybean. Genome **45:** 592–599.

Grant, M., L. Godiard, E. Straube, T. Ashfield, J. Lewald *et al.*, 1995 Structure of the *Arabidopsis RPM*1 gene enabling dual specificity disease resistance. Science **269:** 843–846.

Hamalainen, J. H., K. N. Watanabe, J. P. T. Valkonen, A. Arihara, R. L. Plaisted *et al.*, 1997 Mapping and marker-assisted selection for a gene for extreme resistance to *potato virus Y*. Theor. Appl. Genet. **94:** 192–197.

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

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

Hunst, P. L., and S. A. Tolin, 1982 Isolation and comparison of two strains of *soybean mosaic virus*. Phytopathology **72:** 710–713.

Jeong, S. C., A. J. Hayes, R. M. Biyashev and M. A. Saghai Maroof, 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., L. F. Marek and R. C. Shoemaker, 1996 Resistance gene analogs are conserved and clustered in soybean. Proc. Natl. Acad. Sci. USA **93:** 11746–11750.

Leister, D., A. Ballvora, F. Salamini and C. Gebhardt, 1996 A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. Nat. Genet. **14:** 421–429.

Meyers, B., A. Dickerman, R. Michelmore, R. Pecherer and R. Sivaramakrishnan, 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., and B. Meyers, 1998 Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res. **8:** 1113–1130.

Ori, N., Y. Eshed, I. Paran, G. Presting, D. Aviv *et al.*, 1997 The *I*2C family from the wilt disease resistance locus I2 belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. Plant Cell **9:** 521–532.

Pan, Q., J. Wendel and R. Fluhr, 2000 Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. J. Mol. Evol. **50:** 203–213.

Peñuela, D., D. Danesh and N. D. Young, 2002 Targeted isolation, sequence analysis, and physical mapping of nonTIR NBS-LRR genes in soybean. Theor. Appl. Genet. **104:** 261–272.

Richter, T., and P. Ronald, 2000 The evolution of disease resistance genes. Plant Mol. Biol. **42:** 195–204.

Saitou, N., and M. Nei, 1987 The neighbor joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4:** 406–425.

Shoemaker, R., K. Polzin, J. Labate, J. Specht, E. Brummer *et al.*, 1996 Genome duplication in soybean (*Glycine* subgenus soja). Genetics **144:** 329–338.

Sun, Q., N. Collins, M. Ayliffe, S. Smith, J. Drake *et al.*, 2001 Recombination between paralogues at the rp1 rust resistance locus in maize. Genetics **158:** 423–438.

Swofford, D. L., 2001 *PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods)*, Version 4.0b8. Sinauer Associates, Sunderland, MA.

Wang, Z. X., M. Yano, U. Yamanouchi, M. Iwamoto, L. Monna *et al.*, 1999 The *Pi*b gene for rice blast resistance belongs to the nucleotide binding and leucine-rich repeat class of plant disease resistance genes. Plant J. **19:** 55–64.

Whitham, S., S. P. Dinesh-Kumar, D. Choi, R. Hehl, C. Corr *et al.*, 1994 The product of the tobacco mosaic virus resistance gene *N*: similarity to Toll and the interleukin-1 receptor. Cell **78:** 1101–1115.

Young, N., 2000 The genetic architecture of resistance. Curr. Opin. Plant Biol. **3:** 285–290.

Yu, Y. G., M. A. Saghai Maroof, G. R. Buss, P. J. Maughan and S. A. Tolin, 1994 RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. Phytopathology **84:** 60–64.

Yu, Y. G., G. R. Buss and M. A. Saghai Maroof, 1996 Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. Proc. Natl. Acad. Sci. USA **93:** 11751–11756.

Communicating editor: A. H. Paterson