Theor Appl Genet (2001) 103:406–414

S.C. Jeong · A.J. Hayes · R.M. Biyashev
M.A. Saghai Maroof

Diversity and evolution of a non-TIR-NBS sequence family that clusters to a chromosomal "hotspot" for disease resistance genes in soybean

Received: 25 October 2000 / Accepted: 22 November 2000

Abstract In soybean, genes controlling resistance to numerous diseases have been shown to cluster to regions on several chromosomes. One such vital chromosomal region is on the soybean molecular linkage group (MLG) F flanked by the RFLP markers K644 and B212. Here, genes controlling resistance to bacterial blight, *Phytophthora* root rot, and several viral diseases, as well as QTLs conditioning resistance to corn earworm, root knot nematode, and white mold have been mapped. We have previously identified two classes (b and j) of disease resistance-related nucleotide binding site (NBS) sequences that localize to this cluster of genes. Using both cDNA and genomic analyses, we have studied one multi-gene family of sequences representing the previously reported class j NBS of soybean. This class of NBS resembles the *RPS2*-like NBS sequences. *RPS2* and similar resistance genes are referred to as non-TIR because they do not encode motifs homologous to the Toll-Interleukin-1 region (TIR). By designing PCR primers that specifically target these non-TIR-NBS encoding sequences, we have amplified at least six class j sequence members from soybean. In addition, we have conducted genomic and cDNA library screenings to identify additional class j members. In all, we have characterized 12 class j NBS sequence members. These members have been mapped within a 2-cM region of the soybean F linkage group. We have also identified homoeologous chromosomal regions on linkage groups A2 and E that contain class j NBS sequences. A BLAST search of the GenBank database has shown that non-TIR NBS sequences are present across the legume family. We have compared these non-TIR sequences from other legumes with the soybean clones to assess the level of diversity within this class of disease resistance-related sequences.

Keywords Soybean disease resistance · Nucleotide binding site motif · Gene family

Communicated by G. Wenzel

S.J. Jeong · A.J. Hayes · R.M. Biyashev · M.A. Saghai Maroof (✉)
Department of Crop and Soil Environmental Sciences,
Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061-0404, USA
e-mail: smaroof@vt.edu
Tel.: 1-540-231-9791, Fax: 1-540-231-3431

Introduction

Resistance gene clustering

In soybean, the chromosomal region between RFLP markers K644 and B212 of molecular linkage group (MLG) F is of particular interest for studies of the genetics and evolution of disease resistance genes. Many loci conferring resistance to diverse pathogens have been mapped to this region. These resistance genes include *Rsv1* conferring resistance to soybean mosaic virus (Yu et al. 1994), *Rpg1* to bacterial blight (Ashfield et al. 1998), *Rps3* to *Phytophthora sojae* (Diers et al. 1992), and *Rpv1* to peanut mottle virus (Saghai Maroof, unpublished). In addition, QTLs conditioning resistance to two root knot nematodes, corn earworm, and white mold have been located in this region (Tamulonis et al. 1997a, b; Arahana et al. 1999; Rector et al. 1999).

Clustering of disease resistance genes has been studied in numerous plant species (for a review see Michelmore and Meyers 1998). Recent molecular characterization of these gene clusters has shown that some consist of active and/or inactive variants conditioning multiple resistance specificities for a single pathogenic organism. Examples of multi-allelic clustering of resistance genes to a particular pathogen include the *Cf-9* leaf-mold resistance locus in tomato (Parniske and Jones 1999), the *Rp1* rust resistance locus in maize (Sudupak et al. 1993; Collins et al. 1999), the *Dm* downy mildew resistance loci in lettuce (Meyers et al. 1998a), the *Xa21* bacterial blight resistance locus in rice (Song et al. 1997), and the *M* rust resistance gene locus of flax (Ellis et al. 1995).

Other gene-family clusters consist of multiple genes conditioning resistance to various pathogens. The cluster-

ing of resistance genes to diverse pathogens has been poorly characterized at the molecular level. However, genetic mapping has shown that several examples, such as resistance loci to the fungus *Cladosporium fulvum* and the nematode *Meloidogyne incognita* in tomato (Dickinson et al. 1993), resistance loci for stem rust, scald, and net blotch in barley (Spaner et al. 1998), and resistance loci *Rps2* and *Rmd* in soybean (Polzin et al. 1994) do exist.

Accumulating evidence indicates that the structures of resistance gene clusters are tandem repeats spanning short chromosomal regions. These data have given some insight into the molecular aspects controlling the evolution of new recognition specificities to diverse pathogens. Several hypotheses for the evolution of resistance genes arrayed in complex clusters have been put forward. These include unequal crossing-over (e.g., Sudupak et al. 1993; Song et al. 1997), interlocus recombination (Parniske and Jones 1999), and birth-and-death processes (Michelmore and Meyers 1998). However, evolutionary mechanisms that can be generally applicable to the nature of the clustering of resistance genes remain to be demonstrated.

Cloning of NBS/LRR plant disease resistance genes

Molecular characterization of a number of cloned plant disease resistance genes has provided insight into the conserved sequence structure within this class of plant genes. Most resistance genes cloned encode common conserved motifs characteristic of receptor-like kinases (for a review see Bent 1996). A major class of characterized resistance genes encode a nucleotide binding site (NBS) region and a leucine-rich repeat (LRR) region. The finding of the conserved NBS motif in cloned resistance genes quickly led to PCR amplification of similar sequences from other crops, including soybean and potato, using degenerate primers (Kanazin et al. 1996; Leister et al. 1996; Yu et al. 1996). Most of the NBS sequences have been mapped near known disease resistance loci. Subsequent studies from various plant sources (e.g., Aarts et al. 1998; Collins et al. 1998; Shen et al. 1998; Meyers et al. 1999; Graham et al. 2000) have provided evidence that this approach is useful for genetic and molecular studies of various disease resistance genes.

One of the objectives of the study reported here is to further evaluate the utility of resistance gene-related sequences in soybean for gene discovery. Previously, Yu et al. (1996) identified a superfamily of soybean sequences using degenerate oligonucleotide primers for the NBS region of *N* from tobacco and *RPS*2 from *Arabidopsis*. The soybean clones were categorized into 11 classes: nine classes more closely resembling the *N* gene, and two classes more closely resembling the *RPS2* gene. Studies of additional NBS-LRR disease resistance genes from various plant species have shown that the genes homologous to *N* code for a common N-terminal motif characteristic of the *Drosophila* Toll/Human Interleukin-1 (TIR) genes (Parker et al. 1997). Genes resembling *RPS2* do not encode this TIR motif (non-TIR). The N-terminus of *RPS2*-homologous gene-products do not seem to contain a conserved motif, although several of them have been predicted to contain a putative leucine-zipper motif (Milligan et al. 1998) or, more broadly, a coiled-coil motif (Pan et al. 2000). Relative to the TIR-NBS genes, the non-TIR-NBS sequences probably are not abundant in soybean. With the exception of minor representation of non-TIR sequences in the Yu et al. (1996) study, the presence of non-TIR-NBS in soybean has not been reported in the literature.

Out of 11 classes characterized by Yu et al. (1996), one TIR-NBS, class b, and one non-TIR-NBS, class j, map to the complex region of disease resistance genes on soybean MLG F. It has been demonstrated previously that the class b clone represents a single-copy gene on MLG F (Hayes et al. 2000). In the RFLP analysis of various soybean lines with several restriction endonucleases, the class j clone, NBS 61, hybridizes to about 10–15 restriction fragments. The many cross-hybridizing bands of the class j clone in a genomic Southern led us to believe that it most-likely represents a family of clustered resistance genes in this region. Recently, Hayes and Saghai Maroof (2000) reported a second non-TIR-NBS, R14, obtained from a targeted NBS-based AFLP approach, which also maps to this region on MLG F. R14 and NBS 61 are highly homologous (78%) in a pairwise comparison of DNA sequences, indicating that R14 is another member of the class j NBS sequence family.

Here, we report cloning, mapping, and sequencing of additional class j NBS sequences from this region. We use both PCR and library screening strategies to isolate these clones. The results show that these sequences may represent portions of active resistance genes in this vital cluster of disease resistance loci on MLG F. In addition, we identify and evaluate class j-homologous sequences of related legumes obtained by a BLAST search of the public sequence database.

Materials and methods

Plant materials

The soybean line PI96983 was used as the source of genomic DNA for PCR amplification of class j NBS sequences. Two soybean populations of the parents PI96983×Lee 68 (243 F_2 individuals) (Yu et al. 1996) and V71–370×PI407.162 (149 F_2 individuals) (Maughan et al. 1996) were used to map the genetic location of NBS sequences. The first population was used for mapping clones on the F linkage group, because numerous tightly linked markers have been mapped to this chromosomal region in this population. The second population contains markers spanning all 20 soybean linkage groups and thus was used for mapping loci in regions outside the F linkage group.

Amplification and cloning of class j PCR products

A pair of oligonucleotide primers was designed corresponding to sequences from regions conserved between NBS61 and R14 (two soybean non-TIR-NBS). The primer sequences were: class j-F,

AAT GGG TGG CGT GGG GAA GAC and class j-R, AAT TTA GCG GAG TTA GCA CAG C. Class j-F corresponds to the nucleotide sequence encoding the P-loop amino-acid (M/P)GGVGKT motif conserved between the *N* and *RPS*2 resistance genes. Class j-R corresponds to a region 5′ of the nucleotide sequence encoding the conserved kinase 3a motif that is divergent from other NBS families but highly conserved between NBS61 and R14. This primer set was used for PCR-amplification of NBS sequences of about 300-bp fragment size from the soybean genomic DNA of PI96983. The PCR was in a total volume of 50 µl containing 50 ng of genomic DNA, 1× reaction buffer, 2.5 mM of $MgCl_2$, 320 µM of dNTPs, 1 µM of each primer, and 2U of *Taq* polymerase (Gibco-BRL, Rockville, Md.). The PCR conditions consisted of a hot start (3 min at 94°C) and 34 cycles of 1 min at 94°C, 1 min at 37°C, and 2.5 min at 72°C. The reaction was terminated by a 7-min extension at 72°C, and a 15-min stabilization at 37°C followed by a 4°C soak.

The PCR products of about 300-bp size were purified from agarose gel slices using a Qiaex II gel extraction kit (Qiagen, Valencia, Calif.). The purified products were subcloned into the pCRII-TOPO vector (Invitrogen, Carlsbad, Calif.) or the pCNTR vector (5prime-3prime; Boulder, Colo.) following the manufacturers' instructions.

Screening of genomic and cDNA libraries

Using class j specific probes, we screened both cDNA and genomic libraries of the soybean lines Williams 82 and L81–4420 (a Williams isoline which contains a chromosomal contribution for the resistance gene cluster on MLG-F from PI96983). The Williams 82 cDNA library was obtained from Stratagene (La Jolla, Calif.) and was constructed in the Lambda UniZAP XR vector using 12-day old greenhouse-grown epicotyl tissue. The Williams 82 genomic library was also obtained from Stratagene and was constructed in the FIXII vector. L81–4420 libraries were custom made by Clontech (Palo Alto, Calif.). The cDNA library was made from 10-day old, greenhouse-grown, leaf tissue in the Lambda TriplEx vector. The genomic library was constructed in the EMBL SP6/T7 vector. Libraries were screened according to the manufacturer's protocols.

Reverse transcribed PCR

Total RNA from 2-week old leaf tissue of the line PI96983 was extracted using the TRIzol total RNA isolation reagent from Gibco-BRL (Rockville, Md.). Based on the sequence of class j genomic clones identified in this study, specific primers from the 5′ region were designed. These primers were used to synthesize specific amplified cDNA products using the SuperScript one step RT-PCR kit from Gibco-BRL.

RFLP and sequence analysis

DNA extraction and RFLP analysis were essentially as previously described (Yu et al. 1994) using six restriction endonucleases (*Hin*dIII, *Eco*RI, *Eco*RV, *Xba*I, *Dra*I and *Bam*HI). Mapmaker/exp 3.0 was used for genetic linkage and distance determination (Lander et al. 1987). Marker order was determined at a LOD threshold of 3.0.

DNA sequencing was done using the Dye Terminator cycle sequencing kit and an ABI 377 automated DNA sequencer according to the manufacturer's protocols (Perkin Elmer, Foster City, Calif.). Multiple-sequence alignments were constructed with Lasergene software from DNAstar (Madison, Wis.). PAUP* was used for generating phylogenetic trees (Swofford 2000). Values for non-synonymous and synonymous substitutions (Ka and Ks) were calculated based on the assumptions set forth in Li (1993) and were computed using the FENS v0.9(b1.4) software program which was kindly provided by Dr. Caro-Beth Stewart of the University at Albany, Suny, N.Y..

Results

Isolation of soybean non-TIR-NBS sequences

All of the unique clones identified in this study were designated NTN for non-TIR-NBS followed by a number and a letter 'g', indicating a genomic-derived clone, or 'c', indicating an expressed NBS sequence. PCR-amplification of soybean genomic DNA using class j-F and –R primers resulted in one major product of the expected 300-bp size. This product was gel-purified and cloned. A total of 74 clones were analyzed by RFLP using a set of diagnostic Southern blots containing four soybean DNA samples digested with six restriction endonucleases. Clones were grouped based on hybridization pattern. At least one representative clone from each RFLP pattern-group was sequenced. After sequencing, it was determined that five groups, representing 30 of the 74 clones, contain sequences similar to the non-TIR class j NBS clones. The remaining 44 clones do not resemble NBS-encoding sequences and most likely represent spurious amplification of non-specific products as a result of the low annealing temperature. After sequencing all NBS-encoding clones from the selected hybridization groups, the five clone-types were further classified into seven unique NTN sequence types. Two of the seven sequence types corresponded to the previously reported NBS61 and R14 that were originally used to design the specific primers.

In addition to the PCR-amplified genomic products, we analyzed a number of lambda genomic clones isolated by hybridization of class j-specific probes to the L81–4420 genomic library. From these analyses we identified three additional NTN sequences which were distinct from the PCR-amplified genomic products.

To isolate additional class j NTN clones and to evaluate the expression of the genomic NBS sequences, we screened cDNA libraries from L81–4420 and Williams 82 with several class j probes. Similarly, we used gene-specific primers, designed based on the sequenced genomic clones, to generate reverse transcribed (RT)-PCR products from PI96983. Expression analysis resulted in the identification of four additional NTN sequences. Furthermore, we confirmed the expression of three NTN clones, previously identified by genomic analysis. In all, we have identified 14 unique NTN sequences derived from three soybean lines.

These 14 sequences have been mapped, both by RFLP hybridization and PCR techniques. Two clones, NTN13c and NTN14c (derived from the Williams 82 library), map to loci on MLGs A2 and E, respectively. The remaining 12 have been localized to no less than 14 tightly linked or cosegregating loci on MLG F (Fig. 1). Individual sequence members have been mapped to one locus or several tightly linked loci flanking the *Rsv*1 gene.

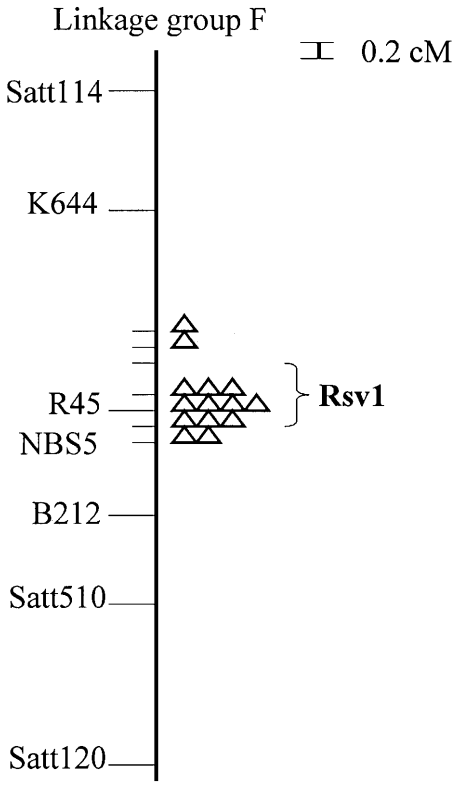


Fig. 1 High-resolution map of the multigene family of class j NBS (NTN) sequences on the molecular linkage group F. All markers were mapped in the F_2 population PI96983×Lee68 which contains 243 F_2 members. *Triangles* represent map locations of individual NTN sequence members. Specific NTN designations were not used because, in some cases, distinguishing the map location of one individual sequence member among several highly related members was not possible

Sequence analysis of the F-linked clones

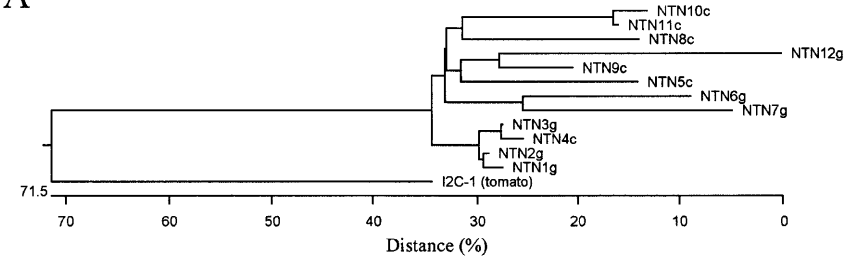
Because of the high-level of homology between NTN members and the inherent complexities this creates for evaluating sequences across genotypes, we only evaluated MLG F NTN sequences derived from the PI96983 source. The NBS-encoding region for the 12 unique F-linked NTN sequences from PI96983 were aligned using the Clustal multiple sequence alignment program. Among the sequence family members, the most divergent members share 77% nucleotide sequence identity and the most related members are greater than 99% identical (Fig. 2B).

One sequence group, including NTN1g, 2g, 3g, and 4c, all show the same unique multi-copy hybridization pattern. This group represents no less than four sequences on the F linkage group. In fact, when used as an RFLP probe in our mapping population, this clone group maps to three unique loci representing four different strongly hybridizing polymorphic bands. These loci all map within a 1-cM region flanking *Rsv*1. The four sequence members differ by only 2–4 nucleotides per member in the NBS encoded region between the conserved p-loop and kinase 3a motifs. The sequence of each PCR-amplified member was confirmed by sequencing multiple representatives.

Another pair of closely related NBS sequences, NTN10c and 11c, differs by only three nucleotides. The genomic clones from which these NBS sequences were derived, were mapped by hybridization, using sequences 3–4 kb downstream from the NBS sequence, in the 3′ untranscribed region. Mapping data using probe sequence from this region were far less complicated

Fig. 2A, B Nucleotide sequence comparison of all 12 class j NBS (NTN) sequences mapping to the MLG F. For a description of NTN clones see Table 1. **A** Phylogenetic tree constructed by the neighbor-joining method (Saitou and Nei 1987) implemented in Lasergene (DNAstar, Madison, Wis.) after Clustal alignment of the 12 sequences, showing relatedness of individual clones. *I2C-1* was used as an outgroup sequence for anchoring the tree. **B** Pairwise sequence distance comparisons of the 12 characterized NTN sequences. The percent identity is calculated by direct comparison without accounting for phylogenetic relationships

A



B

NTN sequence												
1g	2g	3g	4c	5g	6g	7g	8c	9c	10c	11c	12g	
	99.0	98.6	97.9	85.9	84.5	84.9	86.6	85.6	86.9	87.6	81.8	NTN1g
		99.0	98.3	86.3	84.9	85.2	86.6	85.9	86.6	87.3	82.1	NTN2g
			99.3	85.9	84.9	85.6	86.3	85.6	86.3	86.9	81.8	NTN3g
				85.9	84.2	84.9	86.3	84.9	85.6	86.9	81.1	NTN4c
					83.5	82.8	85.9	84.5	84.9	86.3	79.0	NTN5g
						86.6	83.2	82.1	85.9	86.3	78.7	NTN6g
							82.8	81.1	84.5	85.2	77.0	NTN7g
								85.6	87.6	89.0	78.0	NTN8c
									84.5	84.9	79.0	NTN9c
										98.6	79.4	NTN10c
											79.4	NTN11c
												NTN12g

Table 1 Members of soybean class j NBS sequences and their expression

Clone	GenBank accession#	MLG	Source	Expression [a]
NTN1g	AF314546	F	Genomic PCR	ND
NTN2g	AF314547	F	Genomic PCR	ND
NTN3g	AF314548	F	Genomic PCR	ND
NTN4c	AF314549	F	cDNA library/RT PCR	Yes
NTN5c	AF314550	F	Genomic PCR/RT PCR	Yes
NTN6g	AF314551	F	Genomic PCR	ND
NTN7g	AF314552	F	Genomic PCR	ND
NTN8c	AF314553	F	Genomic PCR/RT PCR	Yes
NTN9c	AF314554	F	RT PCR	Yes
NTN10c	AF314555	F	Genomic library/RT PCR	Yes
NTN11c	AF314556	F	Genomic library/RT PCR	Yes
NTN12g	AF314557	F	Genomic library	ND
NTN13c	NS	A	cDNA library	Yes
NTN14c	NS	E	cDNA library	Yes

[a] Expression was detected by RT-PCR or cDNA library screening. ND; not detected. NS; not sequenced

because only a few bands were visualized by Southern hybridization.

Non-TIR-NBS among leguminous species

Using several soybean NBS sequences we have searched the GenBank database of non-redundant nucleotide sequences using BLASTn (Altschul et al. 1997). Of the several hundred NBS sequences which are in the database, there is a clear group of non-TIR-NBS sequences which include all monocot NBS sequences and a small subset of the dicot NBS sequences (Meyers et al. 1999; Pan et al. 2000). Based on our own analysis, one sub-group within the non-TIR sequences is characterized by two conserved tryptophan residues which occur just after the highly conserved kinase-2 region in the deduced amino-acid sequence of these clones (Fig. 3A and B). All of the NTN sequences evaluated in this soybean study fall within this group of sequences.

Among the dicot sequences that fall within this conserved tryptophan sub-group, more than 13 non-soybean, legume sequences, including NBSs from pea, cowpea, pigeon pea and common bean, have been reported in the GenBank database. These sequences were compared with the soybean NTN sequences by Clustal alignment. Interestingly, several of the sequences from other legumes are closely related to one or several class j members (Fig. 3A and B). For example, based on deduced amino-acid sequences, soybean NTN12g is 72% similar to the pigeon pea NBS, AF186634. NTN12g is more similar to this NBS than to any of the other known members of the soybean NTN multigene family. Similarly, a common bean NBS (AF084026) is more related (75%) to the soybean cDNA, NTN4c, than to any of the other reported common bean NBS.

Evolutionary conservation among legume non-TIR sequences

In order to assess the level of conservation within the NBS region of the NTN soybean sequences, we calculated the ratio of non-synonymous (Ka) to synonymous (Ks) substitutions according to the method of Li (1993). This method attempts to assess if there is a selective bias associated with nucleotide changes between sequences. This bias can either be toward maintaining the amino-acid sequence (purifying selection) or toward the development of novel amino-acid sequences (diversifying selection). Values less than one suggest purifying selection and values greater than one indicate diversifying selection. The NBS region of the 12 unique sequences from linkage group F were evaluated for a total of 66 pairwise comparisons. The average Ka/Ks ratio is 0.61; however, only 1/3 of the pairwise comparisons were significantly less than one based on a standard *t*-test. This analysis excludes the highly conserved p-loop and kinase-3a regions because for several of the NTN clones, the primers used to generate the sequences come from these regions.

In order to make a better assessment of the 5′ portion of this gene family, we evaluated six of the genomic and cDNA NTN clones for which we had sequence information across the entire 5′ gene region. Pairwise comparisons were done for the first 1400 bp of these six sequences. This sequence extends from just 3′ of the putative ATG start site to approximately 280 bp 3′ of the highly conserved sequence coding for the GLPLAL hydrophobic region. The average ratio of Ka to Ks is 0.47 with 14 of the 15 pairwise comparisons being significantly less than one at the 0.001 level (Fig. 4A).

In a second analysis, we selected a subset of seven soybean NBS-encoding sequences representing the major subgroups within the class j family. These sequences were compared with six highly similar NBS sequences from other leguminous species including three clones from common bean, two clones from pigeon pea and one clone from cowpea. An analysis of the synonymous to non-synonymous substitutions showed similar results to those seen within the soybean gene family (Fig. 4B). In a

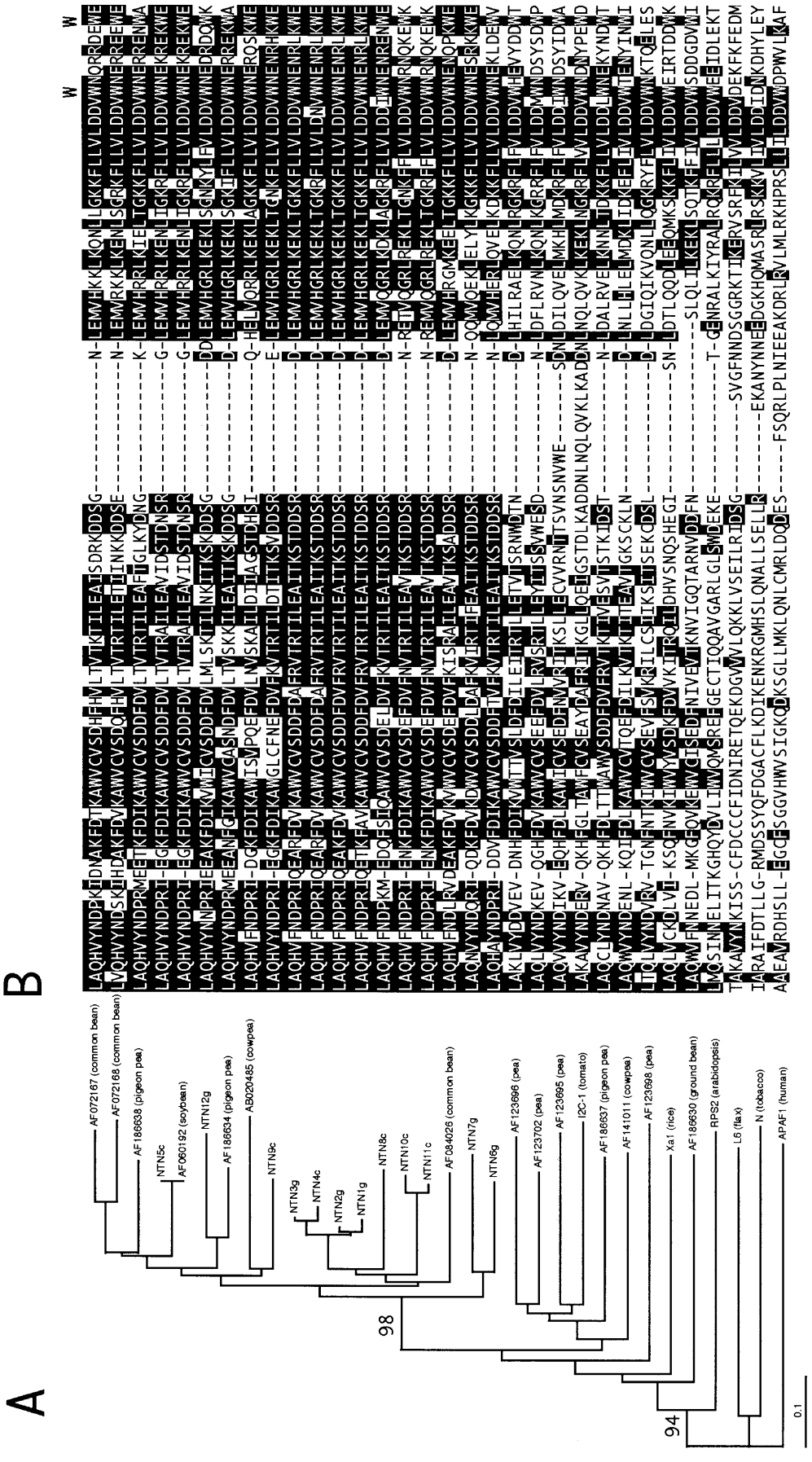


Fig. 3A, B Sequence comparison of the 12 class j NBS (NTN) sequences and closely related non-TIR-NBS from other legumes gleaned from the public sequence database. Nucleotide sequence comparisons were done from 3′ of the p-loop-encoded region to 5′ of the putative kinase 3a-encoded region. The phylogenetic tree (**A**) was constructed after the aligned deduced amino-acid sequences (**B**) based on Clustal alignment followed by manual adjustment. Residues that match the consensus are *shaded black*. NTN sequences in soybean are as described in Table 1. The tree was constructed by the neighbor-joining method (Saitou and Nei 1987) implemented in PAUP* (Swofford 2000). The branch lengths are proportional to the average substitutions per site as indicated by the scale. The soybean and other legume clade and the non-TIR clade were supported by 98% and 94%, respectively, of 100 bootstrap replicates. Other NBS sequences are labelled by their recognized gene name or, in the case of the legume NBS partial coding sequences, by their GenBank accession number. The Ws demarcate two conserved tryptophans which were used to identify this sub-group of non-TIR-NBS. *RPS2* is an example of a non-TIR-NBS resistance gene that does not show both of the conserved tryptophan residues. *N* and *L6* are TIR-NBS-encoding resistance genes. APAF1 is a distantly related NBS-encoding gene from human included as an outgroup to anchor the tree

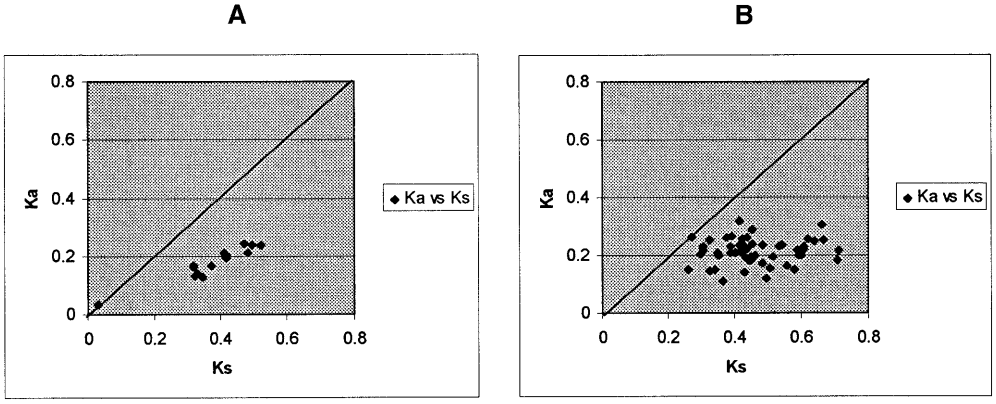


Fig. 4A, B Graphical (scatter plot) display of pairwise comparisons for synonymous substitutions (Ks) vs non-synonymous substitutions (Ka). Equal substitution rates would fall on the center diagonal line. **A** Comparison of six class j sequence members (NTN 4c, 8c, 9c, 10c, 11c and 12 g) across the entire 5′ region of the gene, from just 3′ of the putative start site to about 280 bp 3′ of the highly conserved GLPLAL encoded region. **B** Comparison of cross-species substitution among a sub-group of non-TIR-NBS from soybean, cowpea, pigeon pea and common bean. Calculations were done based on the alignment of NBS encoding sequence from 3′ of the p-loop-encoded region to 5′ of the putative kinase 3a-encoded region

total of 52 cross-species comparisons, the average Ka/Ks was 0.49. Of the 52 comparisons, 27 were significantly less than one at the 0.05 level or higher based on a standard *t*-test.

Discussion

The chromosomal region between the RFLP markers, K644 and B212 of the soybean MLG F, contains numerous members of a multigene family of disease-related sequences. We have isolated and characterized the NBS region for 12 members of this gene cluster belonging to the previously described class j gene family (Yu et al. 1996). Because we had previously identified two related class j sequence members, we were able to design primers to specifically amplify members of this group. While we used several strategies to detect family members, it is likely that we have not identified all of the class j members in this region.

The sequence members evaluated here all share similar RFLP hybridization patterns and significant sequence identity. Several members differ by as few as two nucleotides within the 250-bp stretch between the p-loop and kinase-3 region (see Fig. 2B). Others are significantly divergent from each other and even show in/dels within the stretch. For example, NTN12g and NTN7g are only 77% identical and the former is two codons longer than the latter. It is interesting to note, however, that all 12 sequences contain an open reading frame with conserved sequence motifs across the NBS region.

The level of diversity within the NBS sequence between members of this soybean gene family are similar to those observed for another well-studied gene family surrounding the *Dm*3 locus in lettuce (Meyers et al. 1998b). A pairwise comparison of synonymous to non-synonymous sequences in the lettuce gene family has shown that the 5′ region, including the NBS-encoding portion, is undergoing purifying selection. Our studies show similar levels of structural conservation within this soybean multi-gene family, for the 5′ region. Additionally, we have been able to demonstrate that this selection pattern holds true across several genera within the leguminosae family. Evidence from multiple studies (for a review see Ellis et al. 2000) suggests that the NBS region of NBS/LRR resistance gene products is not involved in pathogen recognition but, instead, may be involved in triggering a resistance response, once pathogen ingress has been detected.

By comparing non-TIR-NBS sequences across several legume species we observed that in several cases there is a stronger relationship between sequence members from related genera than there is between members of the soybean multi-gene family. In other words, NBS sequences from soybean clearly do not form a monophyletic cluster in the phylogenetic tree of NBS sequences from diverse legume species. This observation seems to suggest that this multi-gene cluster may have arisen prior to the divergence of these bean species. In addition, this observation supports the assumption that the non-TIR-NBS gene family in bean species has evolved by the birth-and-death process similar to the evolution of the multigene families of immunoglobin genes and major histocompatibility complex genes (Nei et al. 1997). The birth-and-death process has been suggested for the evolution of plant disease resistance gene clusters by Michelmore and Meyers (1998). However, others have suggested that interlocus recombination or unequal crossing-over may have contributed to the generation and maintenance of multi-resistance gene families in the studies of *Xa21* and the *Cf-9* locus (Song et al. 1997; Parniske and Jones 1999). One common drawback in the evolutionary studies of plant disease resistance clusters, including this study, is that genes or NBS sequences from both functional genes and putative pseudogenes have been compared. Research in plant disease resistance genes is still premature for such an expectation as the distinction between functional genes and pseudo-

genes. Although our present results favor the birth-and-death model for the evolution of this soybean non-TIR-NBS gene cluster, determination of the mechanisms controlling the evolution of bean non-TIR-NBS genes is premature.

The study of this family of NTN sequences from soybean has demonstrated that the characterized members contain an ORF sequence which is undergoing purifying selection. These results suggest a functional nature for the full-length gene products associated with these NBSs. Our future efforts will involve the expansion of the sequence data collected here to obtain full-length gene sequence data. Analysis of the entire gene sequence will provide further insight into the evolution of this important gene family.

Acknowledgements This work was supported by USDA NRICGP, Grant no. 96-35300-3648 and by the United Soybean Board.

References

Aarts MGM, Hekkert BL, Holub EB, Beynon JL, Stiekema WJ, Periera A (1998) Identification of R-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana*. Mol Plant-Microbe Interact 11:251–258

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped Blast and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Arahana VS, Graef GL, Specht JE, Steadman JR (1999) Identification of QTLs for *Sclerotinia* stem rot resistance in soybean using SSR markers. PAG VII P241

Ashfield T, Danzer JR, Held D, Clayton K, Keim P, Saghai Maroof MA, Webb DM, Innes RW (1998) *Rpg*1, 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

Bent AF (1996) Plant disease resistance genes: function meets structure. Plant Cell 8:1757–1771

Collins NC, Webb CA, Seah S, Ellis JG, Hulbert SH, Pryor AJ (1998) The isolation and mapping of disease resistance gene analogs in maize. Mol Plant-Microbe Interact 11:968–978

Collins NC, Drake J, Ayliffe M, Sun Q, Ellis J, Hulbert S, Pryor T (1999) Molecular characterization of the maize *Rp1-D* rust resistance haplotype and its mutants. Plant Cell 11:1365–1376

Dickinson MJ, Jones DA, Jones JDG (1993) Close linkage between the Cf-2/Cf-5 and *Mi* resistance loci in tomato. Mol Plant-Microbe Interact 6:341–347

Diers BW, Mansur L, Imsande J, Shoemaker RC (1992) Mapping phytophthora resistance loci in soybean with restriction fragment length polymorphism markers. Crop Sci 32:377–383

Ellis JG, Lawrence GJ, Finnegan EJ, Anderson PA (1995) Contrasting complexity of two rust resistance loci in flax. Proc Natl Acad Sci USA 92:4185–4188

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

Graham MA, Marek LF, Lohnes D, Cregan P, Shoemaker RC (2000) Expression and genome organization of resistance gene analogs in soybean. Genome 43:86–93

Hayes AJ, Saghai Maroof MA (2000) Targeted resistance gene mapping in soybean using modified AFLPs. Theor Appl Genet 100:1279–1283

Hayes AJ, Yue YG, Saghai Maroof MA (2000) Expression of two soybean resistance gene candidates shows divergence of paralogous single-copy genes. Theor Appl Genet 101:789–795

Kanazin V, Marek LF, Shoemaker RC (1996) Resistance gene analogs are conserved and clustered in soybean. Proc Natl Acad Sci USA 93:11746–11750

Lander ES, Green P, Abrahamson J, Barlow A, Daly JM, Lincoln SE, Newberg L (1987) Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

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

Li W-L (1993) Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J Mol Evol 36:96–99

Maughan PJ, Saghai Maroof MA, Buss GR, Huestis GM (1996) Amplified fragment length polymorphism (AFLP) in soybean: genomic locations, gene action, and evidence of orthologous evolution among three legume species. Theor Appl Genet 93:574–579

Meyers BC, Chin DB, Shen KA, Sivaramakrishnan S, Lavelle DO, Zhang Z, Michelmore RW (1998a) The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. Plant Cell 10:1817–1832

Meyers BC, Shen KA, Rohani P, Gaut BS, Michelmore RW (1998b) Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. Plant Cell 10: 1833–1846

Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnan S, Sobral BW, Young ND (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 RW, Meyers BC (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res 8:1113–1130

Milligan SB, Bodeau J, Yaghoobi J, Kaloshian I, Zabel P, Williamson VM (1998) The root knot nematode resistance gene *Mi* from tomato is a member of the leucine zipper, nucleotide binding, leucine-rich repeat family of plant genes. Plant Cell 10:1307–1319

Nei M, Gu X, Sitnikova T (1997) Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc Natl Acad Sci USA 94:7799–7806

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

Parker JE, Coleman MJ, Szabo V, Frost LN, Schmidt R, van der Biezen EA, Moores T, Dean C, Daniels MJ, Jones JDG (1997) The arabidopsis downy mildew resistance gene *RPP5* shares similarity to the Toll and interleukin-1 receptors with *N* and *L6*. Plant Cell 9:879–894

Parniske M, Jones JDG (1999) Recombination between diverged clusters of the tomato *Cf-9* plant disease resistance gene family. Proc Natl Acad Sci USA 96:5850–5855

Polzin K, Lohnes D, Nickell C, Showmaker RC (1994) Integration of *Rps2*, *Rmd*, and *Rj2* into linkage group J of the soybean molecular map. J Hered 85:300–303

Rector BG, All JN, Parrott WA, Boerma HR (1999) Quantitave trait loci for antixenous resistance to corn earworm in soybean. Crop Sci 39:531–538

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

Shen KA, Meyers BC, Islam-Faridi MN, Chin DB, Stelly DM, Michelmore RW (1998) Resistance gene candidates identified by PCR with degenerate oligonucleotide primers map to clusters of resistance genes in lettuce. Mol Plant-Microbe Interact 11:815–823

Song WY, Pi L-Y, Wang G-L, Gardner J, Holsten T, Ronald PC (1997) Evolution of the rice *Xa21* disease resistance gene family. Plant Cell 9:1279–1287

Spaner D, Shugar LP, Choo TM, Falak I, Briggs KG, Legge WG, Falk DE, Ullrich SE, Tinker NA, Steffenson BJ (1998) Mapping of disease resistance loci in barley on the basis of visual assessment of naturally occurring symptoms. Crop Sci 38: 843–850

Sudupak MA, Bennetzen JL, Hulbert SH (1993) Unequal exchange and meiotic instability of disease-resistance genes in the *Rp1* region of maize. Genetics 133:119–125

Swofford DL (2000) "PAUP:" phylogenetic analysis using parsimony (* and other methods) Ver 4.0b3a. Sinauer, Sunderland, Massachusetts, USA

Tamulonis JP, Luzzi BM, Parrot WA, Hussey RS, Boerma HR (1997a) DNA markers associated with resistance to Javanese root-knot nematode in soybean. Crop Sci 37:783–788

Note added in proof Peñuela et al. have recently submitted a manuscript entitled "Targeted isolation, sequence analysis, and physical characterization of nonTIR NBS-LRR genes in soybean" that details additional study of the class j gene family of soybean.

Tamulonis JP, Luzzi BM, Hussey RS, Parrott WA, Boerma HR (1997b) DNA marker analysis of loci conferring resistance to peanut root-knot nematode in soybean. Theor Appl Genet 95:664–670

Yu YG, Saghai Maroof MA, Buss GR, Maughan PJ, Tolin SA (1994) RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. Phytopathology 84:60–64

Yu YG, Buss GR, Saghai Maroof MA (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