

The Arabidopsis Downy Mildew Resistance Gene *RPP5* Shares Similarity to the Toll and Interleukin-1 Receptors with *N* and *L6*

Jane E. Parker,^{a,1} Mark J. Coleman,^{a,1} Veronique Szabò,^{a,1,2} Louise N. Frost,^a Renate Schmidt,^{b,3} Erik A. van der Biezen,^a Tracey Moores,^a Caroline Dean,^b Michael J. Daniels,^a and Jonathan D. G. Jones^{a,4}

^a Sainsbury Laboratory, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom

^b Cambridge Laboratory, John Innes Centre, Norwich Research Park, Norwich NR4 7UH, United Kingdom

Plant disease resistance genes operate at the earliest steps of pathogen perception. The Arabidopsis *RPP5* gene specifying resistance to the downy mildew pathogen *Peronospora parasitica* was positionally cloned. It encodes a protein that possesses a putative nucleotide binding site and leucine-rich repeats, and its product exhibits striking structural similarity to the plant resistance gene products *N* and *L6*. Like *N* and *L6*, the *RPP5* N-terminal domain resembles the cytoplasmic domains of the *Drosophila* Toll and mammalian interleukin-1 transmembrane receptors. In contrast to *N* and *L6*, which produce predicted truncated products by alternative splicing, *RPP5* appears to express only a single transcript corresponding to the full-length protein. However, a truncated form structurally similar to those of *N* and *L6* is encoded by one or more other members of the *RPP5* gene family that are tightly clustered on chromosome 4. The organization of repeated units within the leucine-rich repeats encoded by the wild-type *RPP5* gene and an *RPP5* mutant allele provides molecular evidence for the heightened capacity of this domain to evolve novel configurations and potentially new disease resistance specificities.

INTRODUCTION

In plants, resistance to disease caused by specialized pathogens often depends on the presence of complementary pairs of genes in the host plant and the pathogen. These are referred to, respectively, as resistance (*R*) genes and avirulence (*avr*) genes, and their expression leads to efficient containment of the pathogen at the site of attempted infection (Flor, 1971; Pryor and Ellis, 1993). The specificity of these interactions suggests that the *R* and *avr* gene products participate in a recognition event that then triggers downstream signaling processes, resulting in the activation of plant defense mechanisms. This is commonly manifested as a local necrotic plant cell reaction called the hypersensitive response. Plant defense-related responses associated with the hypersensitive response include the accumulation of salicylic acid, an oxidative burst, cell wall reinforcements, and the local and systemic induction of genes encoding pathogene-

sis-related proteins (Ryals et al., 1994; Hammond-Kosack and Jones, 1996). However, the precise role of these processes in limiting pathogen growth is unclear.

R genes conferring race-specific resistance have recently been cloned from several crop plant species (Staskawicz et al., 1995; Jones and Jones, 1996; Kunkel, 1996). Analysis of the predicted protein sequences places these genes in distinct classes. The tobacco *N* gene (Whitham et al., 1994), the flax *L6* gene (Lawrence et al., 1995), and the Arabidopsis *RPS2* (Bent et al., 1994; Mindrinos et al., 1994) and *RPM1* (Grant et al., 1995) genes encode proteins containing a nucleotide binding site (NBS) and leucine-rich repeats (LRRs). They mediate resistance, respectively, to viral, fungal, and bacterial pathogens and are all probably localized in the cytoplasm. The tomato *Cf-9* (Jones et al., 1994) and *Cf-2* (Dixon et al., 1996) gene products, specifying resistance to a fungal pathogen, also contain LRRs but are predicted to be predominantly extracytoplasmic with a C-terminal membrane anchor. In contrast, the tomato *Pto* gene (Martin et al., 1993), controlling resistance to a bacterial pathogen, encodes a serine/threonine protein kinase and is therefore quite distinct from the other two classes. However, for its function, it is now known to require an LRR-containing protein, *Prf*, that has sequence similarities with the NBS LRR class of cytoplasmic R proteins (Salmeron et al., 1996). A

¹These authors contributed equally to this work.

²Current address: National Science Foundation Center for Engineering Plants for Resistance against Pathogens (CEPRAP), University of California, 1930 Fifth Street, Davis, CA 95616.

³Current address: Max-Delbrück Laboratory, Carl-von-Linné Weg 10, 50829 Cologne, Germany.

⁴To whom correspondence should be addressed. E-mail jonathan.jones@bbsrc.ac.uk; fax 44-1603-250024.

fourth class is represented by the rice *Xa21* gene for bacterial resistance (Song et al., 1995). This gene encodes a protein that possesses predicted extracellular LRRs, a membrane-spanning domain, and an intracellular serine/threonine protein kinase domain, thus possibly coupling a role in both cell surface recognition and intracellular signaling.

The predicted structural motifs implicate both protein-protein binding, conferred by the LRR domain (Kobe and Deisenhofer, 1994; Jones and Jones 1996), and protein phosphorylation, mediated by a kinase domain (Zhou et al., 1995), in pathogen recognition and activation of defense signals. Also, the limited number of motifs shared among these proteins reinforces the notion that disease resistance in a diverse range of plant species to different pathogen types may operate through similar pathways. A major task in plant pathology now is to understand the precise nature of the *R* gene-mediated recognition events and consequent signaling processes involved.

The interaction between the model crucifer *Arabidopsis* and the biotrophic oomycete pathogen *Peronospora parasitica* is proving to be a particularly attractive system to examine the molecular mechanisms determining disease resistance in plants (Dangl, 1993; Holub and Beynon, 1997). *P. parasitica* is exclusively parasitic on members of the Cruciferae on which it causes downy mildew. It exists in several host-adapted forms, each being specialized to infect a particular plant species or genus (Sherriff and Lucas, 1990; Crute et al., 1994). Characteristic of plant host species that have coevolved naturally with their obligate parasites is an extensive genetic variation between different *Arabidopsis* ecotypes and *P. parasitica* isolates (Holub et al., 1994; Holub and Beynon, 1997). Genetic analyses of these interactions predict the existence of at least 20 different resistance (*RPP*) genes in *Arabidopsis* (Holub and Beynon, 1996). Mapping of several *RPP* loci (Parker et al., 1993; Tör et al., 1994; Holub and Beynon, 1997; Reignault et al., 1996) lends support to the hypothesis that many are clustered in the *Arabidopsis* genome and are likely to have evolved in a manner similar to clustered *R* gene specificities that have been characterized genetically in several crop plants (Hulbert and Michelmore, 1985; Hulbert and Bennetzen, 1991; Jones et al., 1993; Ellis et al., 1995). *Arabidopsis* thus provides an opportunity to isolate naturally polymorphic *R* genes and a system to dissect genetically other components of disease resistance signaling pathways. In particular, examination of the *Arabidopsis*-*P. parasitica* interaction should facilitate analysis of the molecular organization and evolution of *R* genes within potentially complex loci.

We previously characterized the interaction between the *P. parasitica* isolate Noco2 and the resistant *Arabidopsis* ecotype *Landsberg erecta* (*Ler-0*) (Parker et al., 1993). Genetic analysis of segregating populations derived from a cross between *Ler-0* and the susceptible ecotype Columbia (*Col-0*) showed that resistance was conferred by a single locus, designated *RPP5*, which mapped to a 1.2-centimorgan (cM) interval on chromosome 4.

In this study, we describe the positional cloning and characterization of the *RPP5* gene. Sequence analysis showed that it belongs to the class of *R* genes encoding cytoplasmic proteins that possess an NBS and LRRs. The *RPP5* protein exhibits particularly striking N-terminal similarity to the tobacco *N* and the flax *L6* resistance gene products in a region that also resembles the cytoplasmic domains of the *Drosophila* Toll and the mammalian interleukin-1 (IL-1R) receptor proteins (Whitham et al., 1994; Dinesh-Kumar et al., 1995). Examination of the amino acid sequences comprising the LRR regions encoded by *RPP5* and an *RPP5* mutant allele revealed possible mechanisms by which such *R* gene specificities have evolved. Two transcripts are known to be derived from both *N* and *L6* by alternative splicing, giving rise to a predicted full-length and truncated form of each protein. We found no evidence of alternative splicing with *RPP5*. However, we show that a transcript that encodes a protein similar to the *N* and *L6* truncated products is expressed by another member of the *RPP5* gene family, thereby possibly fulfilling a similar but as yet unknown function.

RESULTS

High-Resolution Mapping of *RPP5*

RPP5 was mapped previously using *Ler-0* × *Col-0* recombinant inbred lines (RIs). It mapped to a 1.2-cM interval between the restriction fragment length polymorphism markers m226 (0.9 cM; centromeric with respect to *RPP5*) and g4539 (0.34 cM; telomeric with respect to *RPP5*) on chromosome 4 (Parker et al., 1993). Further definition of recombination events in the RI plants placed *RPP5* telomeric to the restriction fragment length polymorphism marker g13683 (data not shown). A dominant randomly amplified polymorphic DNA (RAPD) marker, OPC18, exhibited complete linkage with *RPP5* in 300 RI plants examined (Parker et al., 1993). The OPC18 primers amplified a 570-bp band from *Col-0* DNA but not from *Ler-0* DNA, and this amplification product was cloned to create a probe, hereafter referred to as C18. C18 was used as an anchor marker to generate a contig of yeast artificial chromosome (YAC) clones spanning the region between the markers g13683 and g4539, as shown in Figure 1A. This defined a physical interval of 280 kb of *Col-0* DNA that contained the *Col-0 rpp5* allele. Further recombination events were identified in this region by using a polymerase chain reaction (PCR)-based strategy with two sets of primers amplifying polymorphic fragments of the locus *SC5* (centromeric to *RPP5*) and g4539 (telomeric to *RPP5*) (see Methods), thus targeting recombinants in a 2-cM interval around *RPP5*. Crossover events between g13683 and g4539 were then positioned in the contig by using a combination of polymorphic YAC end probes and cosmid clones (see Figure 1A). Analysis of crossover events among 1571 F₂ plants narrowed the interval containing *RPP5* to a maximum of 100 kb

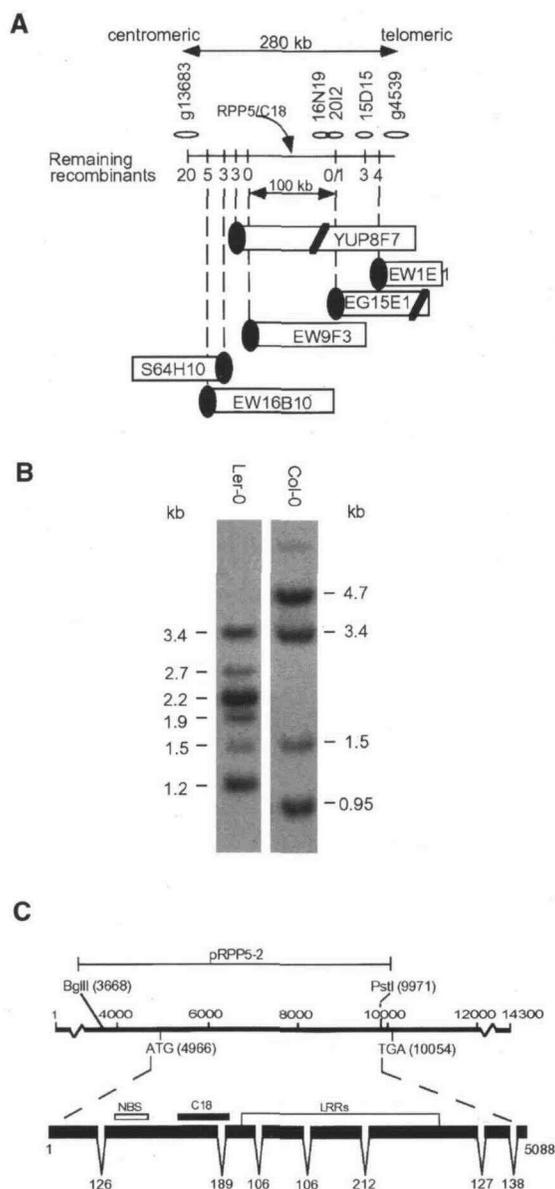


Figure 1. High-Resolution Mapping of *RPP5* on Chromosome 4.

(A) Alignment of YAC clones relative to genetic markers near *RPP5*. Individual YAC clones were positioned by hybridization with the cosmid DNA probes indicated and with the RAPD marker C18 and by cross-hybridization between whole YACs and YAC end probes (black ovals) generated by inverse PCR. YACs were isolated from four libraries (see Methods). Their sizes are drawn to scale. The bold strike-through lines in YACs YUP8F7 and EG15E1 indicate chimerism. Recombination events in selected *Ler-0* × *Col-0* F_2 individuals were placed relative to *RPP5* by probing restriction enzyme-digested genomic DNA with YAC end probes and cosmid markers. The broken lines connect the number of recombinants relative to *RPP5* with the respective probe used. *RPP5* is placed within a 100-kb region contained on YAC clone EW16B10 and cosegregates with C18.

and showed the *Col-0* allele of the locus to be totally contained within the YAC clone EW16B10 (Figure 1A). Partial digestion of EW16B10 with rare cutting restriction enzymes and subsequent probing with YAC end probes revealed that *RPP5* was almost certainly contained within a second smaller YAC, EW9F3 (Figure 1A), because three recombinants were positioned to the telomeric sides of the S64H10 and YUP8F7 YAC ends. These were found to be no more than 14 kb from the EW9F3 centromeric end (data not shown). The analysis demonstrated an uneven distribution of crossover events over the 280 kb of genome examined (Figure 1A). Recombination events were surprisingly rare in the 100-kb region containing *RPP5*.

Analysis of a Small Gene Family Revealed by C18 Hybridization

In all recombinant F_2 plants examined, OPC18 cosegregated with *RPP5* (<0.03 cM) and was physically contained on the 100-kb region defined above (Figure 1A). DNA gel blot analysis of HindIII-digested *Col-0* and *Ler-0* DNA probed with C18 revealed a small polymorphic gene family, as shown in Figure 1B. C18 probing of DNA from all plants that had crossovers within the g13683–g4539 interval (Figure 1A) showed that all polymorphic C18 family members cosegregated with their respective parental *RPP5* genotypes (data not shown). In addition, a comparison of the restriction enzyme-digested YAC clones EW16B10 and EW9F3 with *Col-0* genomic DNA probed with C18 confirmed that all members in *Col-0* were represented on EW16B10 and EW9F3 in a region spanning no more than 50 kb (data not shown). It was considered likely that the gene family occupied a similar interval in the *P. parasitica* Noco2-resistant ecotype, *Ler-0*, although a suitable YAC clone derived from *Ler-0* genomic DNA was not available to confirm this. Because C18-hybridizing DNA represented up to half of the interval of DNA containing

(B) The C18 gene family in *Ler-0* and *Col-0*. HindIII-digested genomic DNA probed with the *Col-0* C18 marker reveals a small, highly polymorphic gene family. The positions and sizes (in kilobases) of DNA molecular weight markers are indicated at left and right.

(C) Sequence features of *Ler-0* insert DNA in binary cosmid clone pCLD29L17. A contiguous sequence of 14.3 kb revealed an open reading frame from nucleotide 4966 (ATG) to nucleotide 10054 (TGA) encoding an NBS/LRR protein. This was shown to be the *RPP5* gene by transformation analysis with a 6.3-kb BglII-PstI subclone of pCLD29L17, designated pRPP5-2. pRPP5-2 encodes an *RPP5* protein lacking the 26 amino acids of exon 8. The *RPP5* coding sequence is expanded, showing the positions and sizes (in nucleotides) of seven introns predicted by reverse transcriptase-PCR (see Figure 3 for further details) and the region with homology to C18. The positions of the motifs comprising the predicted NBS and LRRs are also shown.

RPP5, we selected *Ler-0* binary cosmid clones hybridizing to C18 and transformed representative clones into the Noco2-susceptible Arabidopsis ecotype Nossen (No-0) as a primary screen for *RPP5*-containing DNA.

Isolation of the *RPP5* Gene

Seven *Ler-0* genomic DNA clones in the binary cosmid vector pCLD04541 (see Methods) were identified that hybridized with C18. Stable transformants (T_1 plants) of ecotype No-0 were obtained initially with two cosmid clones, pCLD29L17 and pCLD27E2. Self (T_2) progeny of three independent transformants for pCLD29L17 and two for pCLD27E2 were inoculated with Noco2 and individually tested for kanamycin resistance by using the neomycin phosphotransferase assay. Transformation with pCLD27E2 produced only Noco2-susceptible plants. In contrast, all three pCLD29L17 transformants gave rise to Noco2-resistant T_2 progeny. Cosegregation between the Noco2 resistance phenotype and kanamycin resistance in T_2 and T_3 progeny demonstrated that the *RPP5* gene was contained on pCLD29L17 (data not shown).

Subsequently, two independent transformants of the Noco2-susceptible ecotype Col-0 were obtained with a 6.3-kb BglIII-PstI subclone of pCLD29L17 (denoted pRPP5-2 and shown in Figure 1C) that contained only the putative *RPP5* gene (see below for sequence details). T_2 and T_3 progeny derived from both transformants cosegregated for resistance to Noco2 and the presence of the T-DNA (data not shown).

Resistance to Noco2 conferred by pRPP5-2 in Col-0 was examined in more detail in T_3 transformant lines T-32-4 and T-32-9, which are homozygous for the *RPP5* transgene. Noco2-inoculated seedlings exhibited the same degree of resistance to the pathogen (shown in Figure 2A) as did the natural *RPP5*-expressing ecotype, *Ler-0* (Figure 2B). This can be compared with the fully susceptible response of Col-0 seedlings, on which systemic pathogen growth and profuse asexual sporulation occur (Figure 2C). Microscopic examination of the inoculated material showed that the extent of necrosis at sites of attempted pathogen penetration was the same in Col-0/T-32 transformant seedlings and in wild-type *Ler-0* (data not shown). We concluded from these experiments that the gene on pRPP5-2 is both necessary and sufficient to confer resistance in Col-0 to *P. parasitica* isolate Noco2.

Structure of the *RPP5* Gene

As a first step toward identifying the *RPP5* gene, the nucleotide sequence of the 18-kb *Ler-0* DNA insert of cosmid pCLD29L17 was determined. A shotgun cloning strategy in combination with sequence readings from specific, gap-filling primers were used to obtain 14.3 kb of a double-stranded nucleotide sequence. A candidate *RPP5* gene, encoding a polypeptide with structural features similar to those

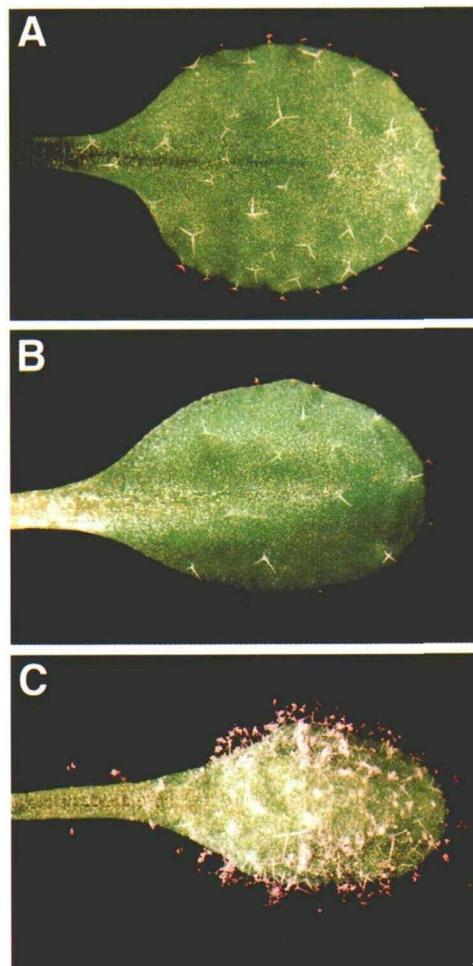


Figure 2. Phenotypes of Different Arabidopsis Lines Inoculated with Noco2.

Nine-day-old seedlings were spray inoculated with a suspension of conidiosporangia. The phenotypes illustrated are of primary leaves 7 days after inoculation.

(A) T-32-9, a Col-0 transformant seedling exhibiting complete resistance to Noco2. This plant is homozygous for the T-DNA and was derived from a Col-0 (T_1) plant that was transformed with the binary cosmid clone pRPP5-2.

(B) *Ler-0*, the Noco2-resistant wild-type control. Resistance is conferred by *RPP5*.

(C) Col-0, the Noco2-susceptible wild-type control permitting profuse Noco2 asexual sporulation and exhibiting typical leaf curling. Susceptibility is the result of lack of *RPP5*.

encoded by several previously characterized plant *R* genes, was identified within the 14.3-kb sequence and was demonstrated by transformation (see above) to be sufficient to confer resistance to Noco2 (Figure 1C). Attempts to isolate a full-length *RPP5* cDNA clone from two *Ler-0* cDNA libraries were unsuccessful. Therefore, reverse transcriptase-PCR

was used to confirm the occurrence of eight exons and seven introns spanning 5088 bp of genomic DNA. Amplification products that unambiguously identified the 3' end of the transcript were not obtained. The derived *RPP5* gene structure is shown in Figure 1C. A sequence with 97% identity to the OPC18 RAPD amplification product forms part of the *RPP5* gene (nucleotides 1520 to 2110 in Figure 3), extending from the end of exon 2, through intron 2, and into part of exon 3 (Figures 1C and 3). Interestingly, subclone pRPP5-2, which was shown to confer full disease resistance on the susceptible ecotype, Col-0 (see above), encodes a protein that is truncated by 26 amino acids, with the clone terminating 3 bp before the start of exon 8 (Figure 3).

The complete *RPP5* nucleotide sequence and derived amino acid sequence are shown in Figure 3. The predicted polypeptide contains 1361 amino acids and has a molecular mass of 154 kD. The sequence was examined for a potential signal sequence and for membrane-spanning regions with the GCG motifs and hydropathy analysis computer programs (Kyte and Doolittle, 1982). The absence of either feature indicated that the protein is probably cytoplasmic. The protein contains a predicted NBS: the sequence QSG-GIGKST (beginning at amino acid residue 216) conforms to the *kinase-1a* (P loop) consensus; the sequences LILLD (beginning at residue 292) and FGSGSR (beginning at residue 314) correspond, respectively, to the *kinase-2* and *kinase-3a* consensus motifs (Traut, 1994). The spacing between these three motifs is consistent with that found in known ATP and GTP binding proteins (Traut, 1994).

The C-terminal region of the RPP5 protein is composed principally of 21 imperfect LRRs, beginning at amino acid residue 576 (Figure 3). These vary in length from 21 to 24 amino acids and contain two regions that have little or no similarity to an LRR consensus. The primary structure of the LRR-containing domain is illustrated in Figure 4A. Figure 4A also shows the two predicted regions of deviation, loop outs 1 and 2, from the canonical LRR sequence. The LRR domain contains regions in which series of amino acids are repeated. For example, LRRs 1, 2, 3, and 4, preceded by part of the non-LRR sequence, are very similar, respectively, to LRRs 5, 6, 7, and 8, preceded by part of loop out 1 (pair 1 in Figure 4A). Three other reiterated sequences are also apparent (Figure 4A). It is notable that the splice junctions of introns 3, 4, and 5 (see Figure 1C) are found at coincident positions in the sequences encoding LRRs 2, 6, and 11 (shown in Figure 4A).

A fragment of *RPP5* genomic DNA that contains all of the coding region with the exception of the last 570 nucleotides was used to probe a gel blot of EcoRI-digested DNA extracted from several different Arabidopsis ecotypes and from the related crucifer crop species *Brassica oleracea* (cabbage), *B. napus* (rape), and *B. campestris* (turnip). The analysis revealed a highly polymorphic family of related genes in the Noco2-resistant ecotypes *Ler-0* (containing *RPP5*), Wasilewskija (*Ws-0*), Praunheim (*Pr-0*), and Oystese (*Oy-0*) and in the susceptible ecotypes Col-0 and Neidersenz (*Nd-0*) (Figure 5A). This re-

sult is consistent with the high degree of polymorphism exhibited by the C18-hybridizing sequences present in *Ler-0* and Col-0, demonstrated earlier (see Figure 1B). Weak hybridization signals were also detected in the Brassica lines by using moderately stringent conditions (Figure 5A), indicating some conservation of *RPP5* sequence in these cultivated species. In contrast, DNA from less-related crop species such as tomato, potato, tobacco, pepper, pea, wheat, barley, and rice did not show any significant hybridization signals with *RPP5* under the same conditions (data not shown).

A Mutant Allele of *RPP5* Contains an Intragenic Duplication of Four LRRs

During the course of this study, *Ler-0* seedlings derived from various mutagenized seed stocks were inoculated with Noco2 and screened for mutations leading to a change from disease resistance to susceptibility. One mutant, designated FL-387, was identified from fast neutron-bombarded *Ler-0* seeds that had a partial loss of resistance. Ten cleaved, amplified polymorphic DNA sequence (CAPS; Konieczny and Ausubel, 1993) or microsatellite (Bell and Ecker, 1994) markers representing the five Arabidopsis chromosomes were used to confirm that FL-387 was not a seed contaminant (results not shown). Genetic analysis revealed that FL-387 contains two unlinked mutations that were separated subsequently by backcrossing to wild-type *Ler-0* (J. Parker, unpublished data). The first mutation, FL-387a, which is described here, contains a structurally altered allele of *RPP5* that was identified by probing FL-387a genomic DNA with an *RPP5* probe (see below). This mutation causes no apparent impairment of the resistance phenotype. The second mutation, FL-387b, confers partial susceptibility to Noco2 and is being analyzed further (J. Parker, unpublished data). Gel blot analysis of restriction enzyme-digested FL-387a genomic DNA, probed with cosmid pCLD29L17, showed an expansion of the *RPP5* C-terminal region containing the LRRs. The apparent insertion caused the disappearance of a 1.2-kb HindIII fragment and gave rise to a novel 1.5-kb fragment, as shown in Figure 5B. Primers were designed to amplify DNA from FL-387a across the region containing the insertion (nucleotides 4041 to 5170 in the wild-type *Ler-0* *RPP5* sequence; Figure 3), and two independently generated 1.3-kb amplification products were cloned and sequenced. This analysis revealed the presence of an in-frame intragenic duplication of 270 bp encoding 90 amino acids, giving rise to four additional LRRs in the open reading frame, as shown in Figure 4B. The duplication is consistent with an unequal crossover event between two regions of 84-bp identity that encodes amino acids in LRRs 13/14 and 17/18 (highlighted in Figures 4A and 4B) of the wild-type *RPP5* protein. Both the gel blot and DNA sequence analysis of FL-387a established that the mutant allele is derived from *RPP5* and not from a seed contaminant.

a. a.	tctgtacaagaactatcataaataagtgacaaccacaaaaatcaatctcta	60	nt.	gagtatctgggtgaactcataatggagtagctgagcttgagaagctgggggtgact	2880
	aaatctcttctgtatgtgtaaaagacttcgaggaagtaagcgaagtttctctc	120		E Y L V E L I M E Y S E L E K L W D G T	
	ctcttctcaactatgctgattactcttcaaaatctctgctctctctctc	180	734	caggtactaattctattagtgataaataatgttagaaaaactaaaaataaaatgttt	2940
1	agtcgctctatctccatcgggctctctctctctctctctctctctctctctc	240	754	aaaatgttcaataacgtgtgtctctctctctctctctctctctctctctctctc	3000
	<u>M A S S S S G R R R Y D V</u>		755	ttctcaagctcagtggtgctgctcgcagcagcttccctcagccatctctcaggg	3060
15	tttttcaagctcagtggtgctgctcgcagcagcttccctcagccatctctcaggg	300	759	l k e m n l r y s n n l k e i p d l s l l	
	<u>P F S F S G V D V R K T F L S H L L K A</u>		779	agccataaacctcaggaatagatctcttttgatggcttctcttggaactctctc	3120
35	ctctgacggcnaatcaatcaatcaatcagatcagatcaggaatcagagaagccacaa	360	799	a i n l e e l d l p f g c v s l v t l p s	
	<u>L D G K S I N T F I D H G I E R S R T I</u>		819	ctcgatcagaatgccactaaactgatctatttagatagatgagcgaatctaga	3180
55	tcgcccctgagcttatatcggcagatagagaagctaggatctcaatcgtctctctcta	420	839	s i q n a t k l l i y l t d m s e c e n l e	
	<u>A P E L I S A I R E A R I S I V I F S K</u>		859	gagttttcaaccctctcaactgaaatctctcagatcctcagatcactgagctgc	3240
75	agaactgctctcaagctggtcttaaatgaattggttgagatcccaagctgctta	480	879	s f p t v f n l k s l e y l d l t g c p	
	<u>N Y A S S T W C L N E L V E I H K C F N</u>		899	gaatttgaaaatctccagcaatcaaaatgggtagctgctgagatagattctcgaac	3300
95	atgatttaggtcaaatggtgattccagtttctcagcagctggtatctcggagtagaa	540	919	n l r n f p a i k m g c a w t r l s r t	
	<u>D L G Q M V I P V F Y D V D P S E V R K</u>		921	aagattgttccggaagggaagatgagatcggtgagaagattgtttctggaacaaga	3360
115	aacagaccgggcaattggaagctcttgaagaacatcggaggtcagcaaggacaac	600	928	r l f p e g r n e i v e d e c f w n k n	
	<u>Q T G E F G K V F E K A T C E V S K D K Q</u>		948	l t c c t c g t g g a g a t a t c t c g a t g c c t t a g a g a t g c c t t g a a t t t c g	3420
135	aacaggggacagaaacaagatggtgcaagctctcagcagatataagcaaatagcg	660	989	t p a g l d y l d c l m r c m p c e a f r	
	<u>P S D Q K Q R W V Q A L T D I A N I A G G</u>		991	ctcagcaaacactcactttctcaatgtagcgctgcaagctgagaagctgagaag	3480
155	agaggtattcttgaagggtagctgattatgattccaatatactgctgctgcttcaaa	720	991	s e q l t f l n v s g c k l e k l w e g	
	<u>F E D L L N G</u>			catccaggtacatggttaatgctatgctgattttgtttaccctctctgtatataaactaa	3540
161	ttgctcagaactatattttgcatagactggttctctttttaggggtgcttcttaat	780		taagatataccaaaattgtttttatggtctgtgtgctgacccaggttatgtcttaccat	3600
	<u>T G C A A A A T T G A C T T T G T T A T A G G C T A A G G C A T A T G T T G A A G A T A C C A</u>	840		atcataataatgtttaataataatttaaacatataagtgataaaaataaaatgatta	3660
173	atgattcttgcataaactctcggcgaaggtttcttgatgactctgctgctgctg	900	921	tcactcagataatgctggaagatccaatggtttttctcagctcgttggaaactcgaag	3720
	<u>D V S N K L I T R S K C F D D F V G G I E</u>			<u>S L G A S L E E</u>	
193	aaactcattatgagcaataaactcagatctgcttggaaactcaggaagctgagaatg	960	928	atggtctctcagaactcgaacactggaagaactccagatctcgaagcccaact	3780
	<u>A H I E A R I S V L C L E S K E A R M V</u>		948	m d l s e s e n l k e a l p d l s k a t n	
213	tcggagctggggcagctcaggttggtagagctaccctcggagagctcttctcagtc	1020	968	ctgaagcttttatgctcagcgggtcgaagcttggtagactctctctcaactctggg	3840
	<u>G I W G Q S G I G K S T I G R A L F S Q</u>		988	l k l l c l s g c k s l v t l p s t i g	
233	aaactctatgacagcttccacatcggcttctcaactataaagaccagctggtatg	1080	988	aatcttcaaaatttgagcgtttgacgttagcagatgcaagctcaggttggattctcgg	3900
	<u>L S S Q F H H R A P L T Y K S T S G S D</u>		1008	m l q n l r l y m n n r c t g l e v l p	
253	acgtctcggatgagctgagcttggtagaagagctctctcggaaatcttaggtcaaa	1140	1028	acaggtgcaactgtcactctcgaacactcagctcagcttctgcaagcttgaga	3960
	<u>V S G M K L S W Q K E L L S E I L G Q K</u>		1048	a c t t t c t g t g t t t c a a t a t t g t a t t d l c t c g c s a a c c g c c t g a a	4020
273	aggaacataaagctagagcattttggttggtaggcaagcttaaatcacaagaagctc	1200	1068	t p p l i s t n i v c c l l y l e n t a i e	
	<u>D I K I E H F G V V E Q R L N H K K V L</u>		1088	gaaatccagactcttcaaaagccagctcagctcagcttctgcaactcaaacctcaaa	4080
293	ttatctctctgtagctgagatatactagagtttcttaagacccttggtaggaagctg	1260	1108	e i p d l s k a t k l e s l l l n n c k	
	<u>L L L D D V D G N L E F L K T L V G K A E</u>		1128	agtttggtagactctctctcaacttggaaactctcaaaatttgagagctgtacatg	4140
313	aatggtttggatctggaagcaatggtgctcactcaagatggcaactctcagga	1320	1148	s l v t l p s t i g n l q n l r r l y m	
	<u>W F G S G S R I V I V T Q D R Q L L K A</u>		1168	aacagatgcaagggctgagcttctcagcagatgctcagcagctcagcagctcgaac	4200
333	ctcactgagatgactctatagagtgagctgacatcgaagctctgctcttaaga	1380	1188	n r c t g l e l e l p t d v n l s s l e t	
	<u>H E I D L V Y E V K L P S Q G L A L K M</u>		1208	ctcagctcagctggttctcaagtttgagaactctctctcagctcagctcagctcga	4260
353	tgatctcccaatagcttttgggaagactctccactcagatgattttaaagcaactagc	1440	1228	l d l s g c s s l r t f l p l i s t r i e	
	<u>I S Q Y A F G K D S P F D D F K E L A F</u>		1248	t g t c t a t c t a g a a a c a c c g c a a t t a g a a g t t c c c t g g a t t a g g a t t c a c g	4320
373	ttgaagtggcagagctctcgttctctctctctctctctctctctctctctc	1500	1268	c l y l e n t a i e i e v p c c i e d f t	
	<u>E V A E L V G S L P L G L S V L G S S T L</u>		1288	t g t c a c t g a c t a c g a a c c a a c t a g a a g t t c c c t g g a t t a g g a t t c a c g	4380
393	taaaaggaagggcaagatgagtggtgagatgctcagctcagaaatctcgaatctcag	1560	1306	a g g t c a c t g a c t a c g a t g a t t g t t g c a g a a g t t g a a a a c t c c c a a c a t t	4440
	<u>K G R D K D E W V K M M P R L R N D S D</u>		1326	r l t v l r m y c c q r l k n i s p n i	
413	atgataaactgaggaacactcaagatcggctcagatgaggttaaaataaanaaactag	1620	1346	t c a g a c t a g a t t c a c g t c g c a g a t t c a g a c t g a g a g t g t c a a g g c f	4440
	<u>D K I E E T L R V G I D R L L N K K N R E</u>		1366	f r l t s l t l a d f t d c r g v i k a	
433	agttatttaagctcagctgatttttcaactggttttaaagctcagtaactcaagaat	1680	1386	t t g a t g a c a c t g g t g g t a c a t t g a a g a c t t c t c g g a t t a c c a t a t c t	4500
	<u>L F K C I A C F P N G F K V S N V K E L</u>		1406	g a a a a c t g a a t a t a c a t g t a a g c t t c t c g a a g c t t c t g a t t a t t a c t g a t	4560
453	taacttgaagatgagctgggctcaactggtgctgagagctccctcactagatataac	1740	1426	e n i e y t c e r f w d a c s d y y s d	
	<u>L E D D V G L T M L A E K S L I R I T P</u>		1446	g a c t t t g a g g t a a a t c g g a a c c a a t t a g a t t g c a a c g a t g a t g c a a g t t g g a d	4620
473	cggttggatatacagatgcaacatttgcagagaaatgggttagagaatgtagctgctg	1800	1466	d f e v n r n p i r l s t m t v n d v e	
	<u>G G Y I E M H N L L E K L G R E I D R A</u>		1486	t t a a g t t t g t g c c a t t a c g a t a a g a a g c t t c a c g c t t g t a t t c t a t	4680
493	caaagtccaagggtaactcctggaacactcaatctcagcagaaatcttgagatatacag	1860	1506	f k f c c s i t i k e c g v r l l y v y	
	<u>K S K G N P G K R Q F L T N F E D I R E</u>		1526	c a a g a a c a g a c a a c a a c a a c a a c t c a g a a a c c a g a a g c g g t a a c c t t	4740
513	aagatgacaggaacactgtagtatttttgcagatccttcaaaccttgaatgtagac	1920	1546	g e t e h n q q t t r s k k r m r	
	<u>V L T E K T</u>		1566	t g c c a a a t t a g a g t a a g c a a c t t t t g a c t g a t t g t t t a t a g g a	4800
533	tttatcaataaactgtaattttgggagtgataaaacttaagcaattgttgcacctag	1980	1586	t c a a a t a c c a t a g c a c a g a c t a t t t g a t a g a t c a t g t t g a t a t a a t g c a g a t	4860
	<u>A T A T G T T C T C T G G A T T T T T C A G G G A C C G A A C T C T C T T G G A A T G T G C</u>	2100	1606	g a c a t c g g g a c a t t c g a a g a a t a c a a c t a c c t a t g g c a a t t g a g g a c a c	4920
553	caacaccggatatactcagcaaggtctctcaatagatgaaactcctcaagagca	2160	1626	t s g t s e e d i n l p y g q i v a d t	
	<u>H P Y L T T R S F L I D E K S F K G M</u>		1646	a g g a t g g c g c t t a a a c a g a g c t t c g t t a g g c a g g a a g a a c a t c a a c	4980
573	tcgtaatactcaactatcagaaatggttattgctcagatgggttctcactcagagc	2220	1666	g l a a l n t e l s l g q g e a s s t t	
	<u>R N L Q Y L E I G Y W S D G V L P Q S L</u>		1686	a t t c t a g a g g g g a a g c t t g t g g t g a t a c a t g a t a a a g a a c a a g a t g a	5040
593	tcgcttattctctgtaaaccaaaagctatggtggatatacttccatbgaagcgtt	2280	1706	s l e g e a l c v d d y m i n e o q d e	
	<u>V Y F P R K L R L W W D N C P L K R L</u>		1726	a c a a t a c c t a t c t g t a t c c t g t t a t g g t a a c t g a a g a c t t t a t c a t c t g t t t g	5100
613	tcgcttcaatatttaagctgagatctggttgaactcagaatggtgaatagtagaactg	2340	1746	c t c t t t t t a g g a t a a c t g g g a c c a t t a t a a a t t a a t a t g a t a a t g a a a	5160
	<u>P S N F K A E V L E R L M V N S K L E</u>		1766	a g a t t c a t a g g t t t g a c t t t g a c a a c a a g c a a t t t t t g a c a t a t a g a c g a t	5220
633	agaagctggtggatggaactcaggtactaattttttagatgcaatttotaacaataaa	2400	1786	g a a a t c c a g a t c t t c t t a g c c a a a a c c t c g a g g a a t a a a c t t g a a g a a t c g a a	5280
	<u>K L W D G T Q</u>		1806	e i p d l s l a i n l e e l n l e e c e	
653	aactaaaaataaaatgtttaaattgtcattacgtgtgctctctcttccctatatt	2460	1826	t c t t t g g a c a c t c t c t c a g t c a g a a t g c c a t t a a c t g a g g a g t t a a t t g t	5340
	<u>T G T T T C A G C C C T T G A A G T C T C A A G A A G A T G G A T T G T A T A A T C T C A A A T G A A</u>	2520	1846	s l e t l p s s i q n a i k l r e l n c	
673	gaaatccagatcttctttagccataaacctcggaggaatcaactcggagaaatcgcaa	2580	1866	t g g g g g g c t a t a a t a g a t t a a a a c a t a g a a g c a t g t a a t c t g a a t a t c t a	2700
	<u>E I P D L S L A I N L E E L N L E E C E</u>		1886	w g g l l d l d l k s l e g m c n l e y l	
693	tccttggagacactctctcagctcagaaatgcaatgcaatgcaatgcaatgcaatg	2640	1906	t c a g t t c a g t t g g t c a a g t a g g a a c c a c t c a g g c a t c g t t a t t c c c t g t a a s	2760
	<u>S L E T L P S S I Q N A I K L R E L N C</u>		1926	s v p s w s s r e c t q g i v y f p r k	
713	tcggggggcttataatagatttaaaactcagaaagctgcaatgcaatgcaatgcaatg	2700	1946	c t c a a a a g t a t t g t g g a a a t g t c a a t g a a g c g t t g c c t c t a a t t t a a g g c t	2820
	<u>W G G L L I D L K S L E G M C N L E Y L</u>			<u>L K S V L W T N C P L K R L P S N F K A</u>	

Figure 3. Nucleotide Sequence of the *RPP5* Gene and the Predicted Amino Acid Sequence.

The sequence shown includes 200 nucleotides (nt.) 5' to the start site (ATG) and 351 nucleotides 3' to the stop codon (TGA). The positions of the seven introns were confirmed by reverse transcriptase-PCR analysis. The *kinase-1a* (P loop) domain (GQSGIGKST), *kinase-2* domain (LILLD), and *kinase-3a* domain (FGSGSR) comprise a predicted NBS within exon 2 and are underlined. The asterisk marks the stop codon. A consensus polyadenylation sequence within the untranslated 3' region is also underlined. Nucleotides shown in boldface (nucleotides 145 to 1722) represent the sequence of identity in transcripts of *RPP^{tr}*, a truncated form of *RPP5* encoded by one or more other members of the *RPP5* gene family (see also Figures 6B and 6C). The *RPP5* sequence GenBank accession number is U97106.

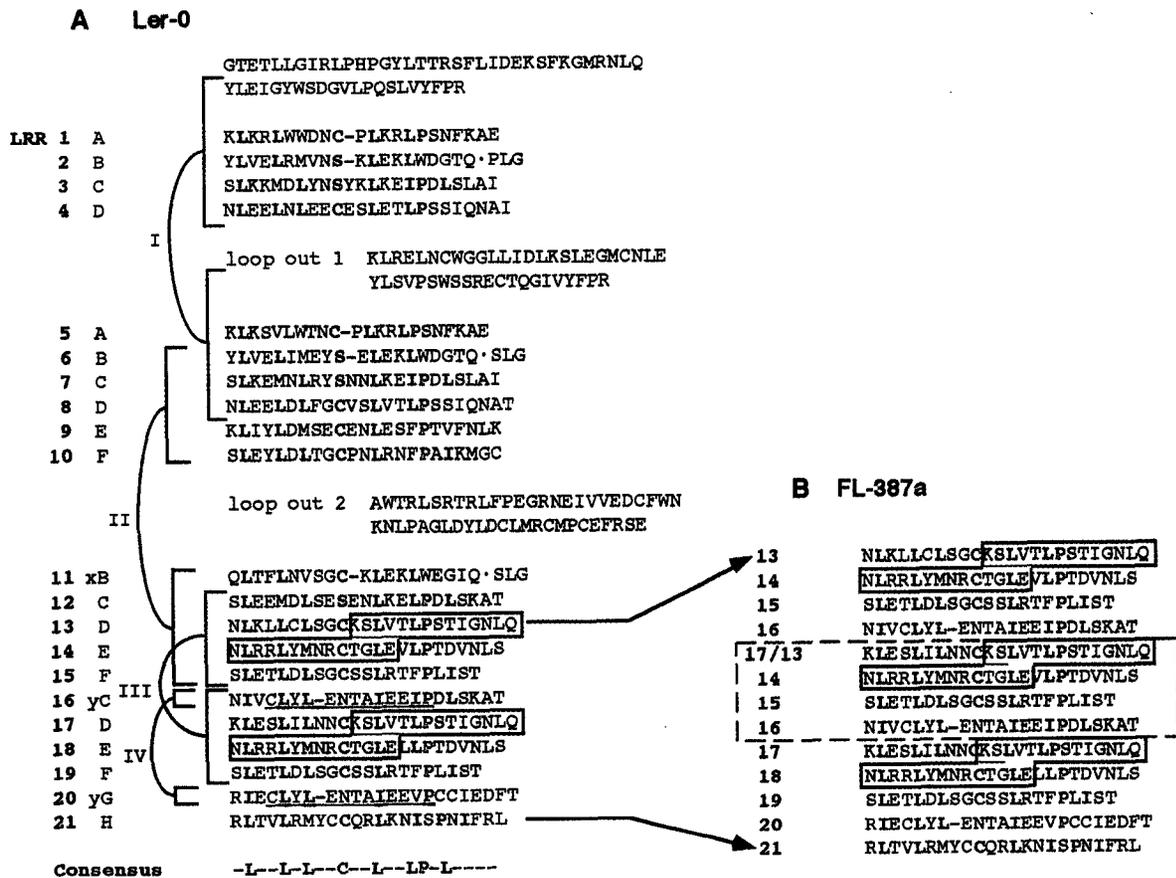


Figure 4. Primary Structure of the Predicted LRRs of *RPP5* in Wild-Type *Ler-0* and FL-387a Plants.

(A) Wild-type *Ler-0* plants. The conserved L or I residues of each of the 21 LRRs are highlighted in blue. Other conserved amino acids (S, C, and P) are shown in red. Two regions that do not conform to the canonical LRR structure are shown as loop outs 1 and 2. LRR 5 conforms weakly to the LRR consensus and so alternatively could be considered as a continuation of loop out 1. Four series of amino acids containing highly similar sequences are shown (bracket pairs I, II, III, and IV). The reiteration of LRR types A, B, C, D, E, and F (in LRRs 5 to 10) is represented in other LRRs with the corresponding letters A to F. Only the latter parts of LRRs xB (11) and yC (16) contain repeated sequences. Also, the sequence similarity (underlined) of series IV is contained within otherwise divergent repeats yC (16) and yG (20). The occurrence of intron-exon splice junctions conserved between introns 3, 4, and 5 is indicated with a dot. Two stretches of 28 amino acids representing 84-nucleotide sequence identity between LRRs 13/14 and 17/18 are highlighted (boxes with solid lines), indicating possible substrates for unequal crossing-over that would produce the duplication observed in FL-387a.

(B) Mutant FL-387a plants. In the *RPP5* sequence of FL-387a, LRRs 13 to 16 are repeated, resulting in a 90-amino acid duplication (box with a broken line). An increase from two to three copies of the 28 amino acid repeats in LRRs 13/14 and 17/18 (boxes with solid lines) is consistent with an unequal exchange event, giving rise to the duplication.

Analysis of the *RPP5* Gene Transcript

A gel blot containing poly(A)⁺ RNA from *Ler-0* and mutant FL-387a seedlings was probed with a 776-bp *RPP5* C-terminal fragment. This revealed a transcript of ~4 kb (Figure 5C, lane 1), corresponding to the size expected for a full-length mRNA encoding 1361 amino acids. The weakness of the hybridization signal, however, indicated that *RPP5* mRNA is present at very low levels, at least in healthy *Ler-0* seedlings.

The *RPP5* allele in FL-387a possesses an in-frame duplication of 270 bp within the coding region (see above) and produced a correspondingly higher molecular weight transcript (Figure 5C, lane 2). This provided confirmation that the 4-kb transcript observed in wild-type *Ler-0* plants is indeed encoded by the *RPP5* gene. A very weak hybridization signal of ~1.8 kb was also apparent in *Ler-0* plants. This may have been the result of hybridization with a related gene transcript (see below).

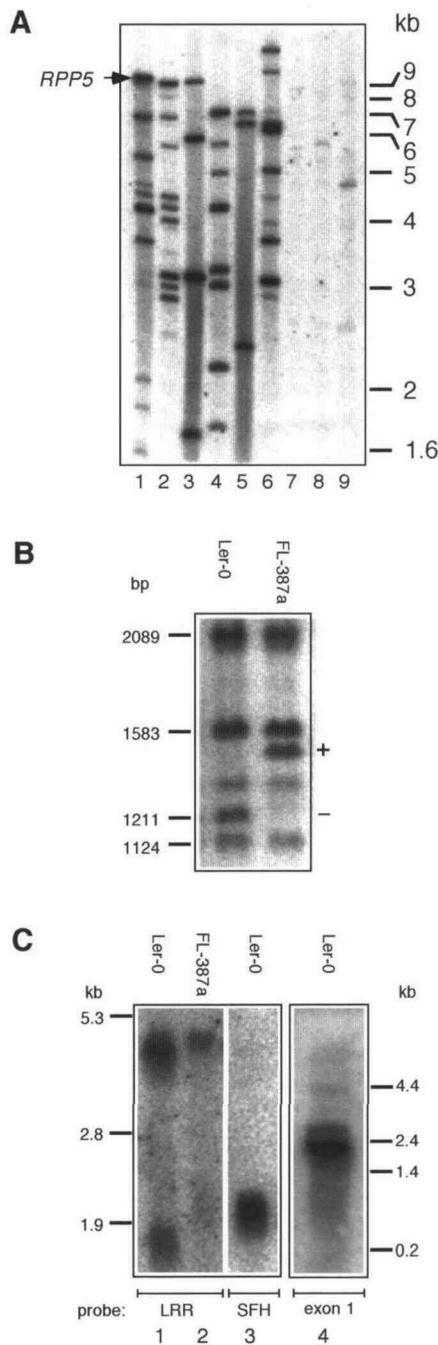


Figure 5. DNA and RNA Gel Blot Analysis of *RPP5*.

(A) DNA gel blot analysis of EcoRI-digested DNA (1 μ g per lane) extracted from different *Arabidopsis* ecotypes and *Brassica* species, probed at medium stringency with the *Ler-0 RPP5* gene (nucleotides 238 to 4308 in Figure 3), reveals a small, polymorphic gene family. *RPP5* gene members are present in Noco2-resistant (*Ler-0*, lane 1; *Ws-0*, lane 4; *Pr-0*, lane 5; *Oy-0*, lane 6) and susceptible (*Col-0*, lane 2; *Nd-0*, lane 3) *Arabidopsis* ecotypes. Weak hybridization signals can also be seen with *B. oleracea* (lane 7), *B. napus* (lane 8), and *B. campestris* (lane 9). Four times more DNA was loaded in lanes 7 to 9.

RPP5 Shares N-Terminal Similarity to the Toll and IL-1R Receptors with N and L6

Analysis of amino acid sequence alignments between the predicted *RPP5* gene product and GenBank database sequences showed that it is most similar to proteins encoded by the tobacco *N* and flax *L6* resistance genes. Notable is the high level of sequence identity in exons 1 and 2 and the similar positions of the intron–exon splice junctions of the full-length transcripts, giving rise to exons 1, 2, and 3 (Figure 6A). A computer alignment of amino acids in *RPP5*, *N*, and *L6* exons 1 is shown in Figure 7A, together with similar sequences from the cytoplasmic domains of Toll and IL-1R. The latter similarity was noted previously by Whitham et al. (1994) for the *N* gene product. Corresponding secondary structure predictions were made for the five aligned sequences, as shown in Figure 7B. The analysis reveals a similar occurrence and spacing of alternating β -strand elements and α -helices, although the predicted IL-1R secondary structure deviates over a region otherwise structurally conserved between Toll and the three plant R proteins (highlighted in Figure 7B).

A Truncated Form of *RPP5* Is Encoded by a Member of the *RPP5* Gene Family

Both *N* and *L6* produce transcripts that encode both a full-length and truncated protein (Whitham et al., 1994; Lawrence et al., 1995). The transcript encoding truncated N (N^{tr}) arises

The 9.6-kb EcoRI fragment in *Ler-0* representing *RPP5* is marked by an arrow. The positions and lengths (in kilobases) of DNA molecular weight markers are indicated at right.

(B) Gel blot analysis of HindIII-digested DNA extracted from wild-type *Ler-0* and FL-387a plants probed with cosmid pCLD29L17. A novel 1583-bp band is present (+) in FL-387a, with the corresponding loss of a 1211-bp band (–), revealing an insertion of \sim 300 bp. The sizes of the *RPP5* HindIII digestion fragments in base pairs are indicated at left.

(C) RNA gel blot analysis of 8 μ g of poly(A)⁺ RNA isolated from wild-type *Ler-0* and FL-387a plants (lanes 1 and 2) probed under high-stringency conditions with a C-terminal *RPP5* gene probe (nucleotides 4041 to 4817 in Figure 3). Exposure to a plate used with a PhosphorImager (Fuji Photo Film Co., London, UK) was for 4 days. The major *RPP5* transcript of \sim 4 kb in *Ler-0* is shifted to \sim 4.3 kb in FL-387a. The same RNA gel blot (lane 3; only *Ler-0* is shown here) was reprobed with a 389-nucleotide sequence representing the SFH present in one or more members of the *Ler-0 RPP5* gene family (see also Figure 6). This reveals a single hybridization signal of \sim 1.9 kb. A different RNA gel blot (lane 4) was probed with an N-terminal sequence (560 nucleotides) representing exon 1 of *RPP5*. A strong signal at \sim 1.9 kb and a very weak signal at \sim 4 kb are shown. The positions and lengths (in kilobases) of RNA molecular weight markers are indicated at left and right.

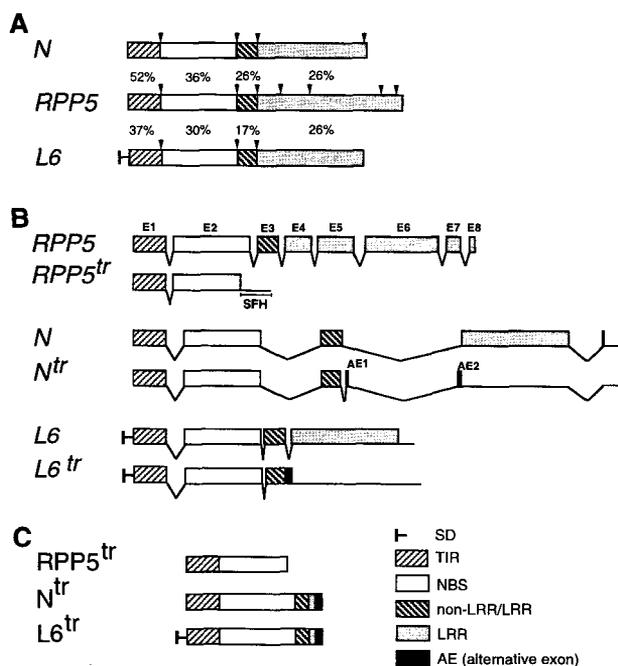


Figure 6. Comparison of the Major Structural Features of *RPP5*, *N*, and *L6*.

(A) The predicted amino acid sequences of the *RPP5*, *N*, and *L6* full-length proteins show extensive similarity. Intron positions are marked by arrowheads. Percentage of amino acid identity between *RPP5*, *N*, and *L6* is indicated for four domains, marked by different shading: domain 1 (exon 1, denoted by TIR) has similarity with the cytoplasmic domains of Toll and IL-1R (see also Figure 7); domain 2 (exon 2, NBS) contains motifs that constitute a nucleotide binding site; domain 3 (exon 3, non-LRR and LRR) contains the start of the LRRs; domain 4 (exons 4 to 6 in *RPP5* and LRR) corresponds to the C-terminal part of the protein containing the remaining LRRs. The same analysis between *RPP5* and *RPS2* reveals 14, 23, 23, and 29% amino acid identity, respectively, in domains 1 to 4. A signal domain (SD) is predicted only for *L6*.

(B) Generation of alternative transcripts from *RPP5*, *N*, and *L6*, giving rise to similarly truncated proteins by three different mechanisms. In contrast to the full-length *RPP5* transcript (E1 to E5 represent exons 1 to 5), a shorter transcript encoding *RPP5* exon 1 and most of exon 2, followed by the SFH sequence, is expressed by one or more members of the *RPP5* gene family. The *N* and *L6* genes produce both full-length and alternatively spliced transcripts (see Whitham et al. [1994], Lawrence et al. [1995], and text for details). AE1 and AE2 are alternative exons of *N*^{tr}.

(C) The truncated *RPP5*, *N*, and *L6* proteins possess the TIR and NBS domains but are essentially devoid of LRRs.

from alternative splicing of a 70-bp exon from within intron 3 (Figure 6B). The transcript encoding truncated *L6* (*L6*^{tr}) is produced when intron 3 remains unspliced and there is readthrough to a stop codon six amino acids beyond the normal intron 3–exon 4 junction (Figure 6B). The high degree of

sequence similarity between these two *R* genes and *RPP5* raised the possibility that *RPP5* may also undergo alternative splicing. We adopted a reverse transcriptase–PCR strategy essentially identical to that used for *L6* to identify incomplete processing of the *RPP5* transcript (see Methods). A single product was detected, and this corresponded to normal splicing of intron 3 (data not shown). Therefore, incomplete processing events similar to those for *L6* were not detected for *RPP5*.

Two *Ler-0* cDNA libraries were screened with an *RPP5* probe to identify cDNAs encoding *RPP5*. Three cDNA clones (clones 9, 11, and 35) were obtained that had long poly(A) tails (>30 nucleotides) and were therefore judged not to have resulted from internal priming events. Clones 9 and 11 represented 5' truncated forms of clone 35 (data not shown). The nucleotide sequence of clone 35 is 1882 bp long, and its first 1450 bp are identical with the 5' end of *RPP5*. The homology extends from 54 bp 5' to the probable ATG and includes exon 1 (479 bp) and the first 917 bp of exon 2, as shown in Figures 3 and 6B. The homology with *RPP5* breaks down abruptly 158 bp 5' to the intron 2 splice donor site (Figure 6B). The 3' end of clone 35 (389 bp, excluding the poly[A] sequence) has no homology with *RPP5*, and its sequence has been designated the SFH (for sequence from homology; see below). SFH DNA was used in hybridization experiments to probe genomic DNA and cosmid clones known to contain *RPP5* and several other members of the *RPP5* gene family. Results showed that the SFH is not present in cosmid pCLD29L17, containing *RPP5*, but is represented in cosmids that contain other members of the gene family. *Ler-0* apparently contains four copies of the SFH, all located at the *RPP5* locus (data not shown). Thus, the SFH-containing transcript is not derived from *RPP5* but from one or more members of the *RPP5* gene family. The transcript encodes a 466–amino acid truncated form of *RPP5* that we denote *RPP5*^{tr}. The first 465 amino acids of this protein are identical with the first 465 amino acids of *RPP5*. In *RPP5*^{tr}, amino acid 466, a tyrosine, is the first amino acid encoded by the SFH. The tyrosine codon is followed by a stop codon. A comparison of the *RPP5*^{tr}, *N*^{tr}, and *L6*^{tr} products reveals them to be structurally very similar, with each encoding an NBS-containing protein that is essentially devoid of LRRs (Figure 6C).

An SFH DNA probe detected a hybridizing band of ~1.9 kb on an RNA gel blot (Figure 5C, lane 3), indicating that clone 35 represents the complete *RPP5*^{tr} transcript. To assess the relative levels of the *RPP5* and *RPP5*^{tr} transcripts, a second RNA blot was probed with an *RPP5* N-terminal probe, encoding only exon 1. A weakly hybridizing band at 4 kb corresponding to the predicted full-length *RPP5* transcript and a much stronger band at 1.9 kb, coincident with the SFH hybridization signal, were observed (Figure 5C, lane 4). Therefore, the *RPP5*^{tr} transcript appears to be more abundant than the *RPP5* mRNA.

In summary, a truncated form of *RPP5* is expressed in *Ler-0* plants that is structurally similar to the alternative

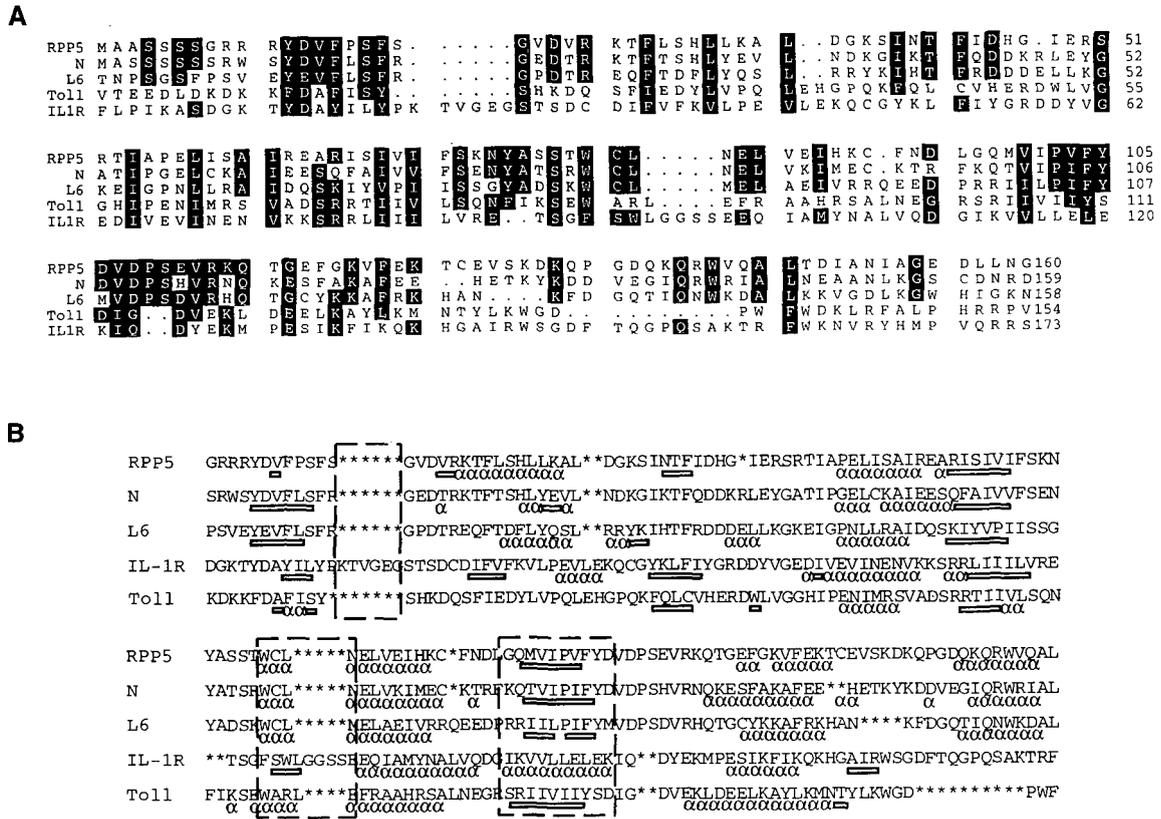


Figure 7. Primary and Secondary Structure Alignments between RPP5, N, L6, IL-1R, and Toll.

(A) Prettybox representation of a Pileup analysis (Genetics Computer Group [Madison, WI] package; see Methods) of the N-terminal regions of RPP5, N, and L6 and part of the cytoplasmic domains of Toll and IL-1R. Shown are RPP5 amino acids 1 to 160 (exon 1), N amino acids 1 to 159 (Whitham et al., 1994), L6 amino acids 51 to 158 (minus the putative signal extension; Lawrence et al., 1995), IL-1R amino acids 352 to 527 (Sims et al., 1989), and Toll amino acids 848 to 999 (Hashimoto et al., 1988). Identical and similar residues are in black boxes. Sequence gaps are indicated with dots.

(B) Comparison of the secondary structure predictions for RPP5, N, L6, IL-1R, and Toll derived from the aligned sequences shown in (A). The alignments extend from RPP5 amino acids 8 to 146 of exon 1. Potential α -helix (α) or β -strand elements (open rectangles) are shown beneath each sequence. Three regions that show deviation of the IL-1R sequence from the other sequences are highlighted by boxes with broken lines. The asterisks indicate gaps accommodated in the sequence alignments.

splice products of *N* and *L6*. However, it is derived from one or several members of the *RPP5* gene family and not from *RPP5* itself.

DISCUSSION

Here, we report the positional cloning of *RPP5*, a gene in Arabidopsis ecotype Ler-0 that specifies resistance to the *P. parasitica* isolate Noco2. The gene is sufficient to confer full resistance to Noco2 when transferred to the genetically susceptible ecotype Col-0.

Comparison of the predicted *RPP5* amino acid sequence with GenBank database sequences shows that it belongs

to the NBS/LRR class of plant *R* genes. To date, this class comprises the tobacco *N* gene mediating viral resistance (Whitham et al., 1994), Arabidopsis genes *RPS2* (Bent et al., 1994; Mindrinis et al., 1994) and *RPM1* (Grant et al., 1995) mediating bacterial resistance, and the flax *L6* gene specifying rust resistance (Lawrence et al., 1995). The predicted protein products of all five genes are of a similar size (909 to 1361 amino acids) and exhibit a similar structural organization. All encode proteins with large C-terminal domains composed principally of LRRs and, in the N-terminal portion, motifs that constitute an NBS. These *R* genes are predicted to encode cytoplasmic proteins, although the *L6* gene product possesses a potential signal peptide, indicating that it may enter a secretory pathway (Lawrence et al., 1995). The presence of the same conserved domains in

RPP5, specifying resistance to an oomycete pathogen, strongly reinforces the hypothesis that common recognition and resistance signaling mechanisms operate in different plant species to combat infection by taxonomically unrelated pathogens. This conserved mode of action is also likely to require binding of either GTP or ATP.

Surprisingly, the C-terminal 26 amino acids comprising exon 8 do not appear to be required for *RPP5* gene function, because a subclone lacking this coding region conferred on the genetically susceptible ecotype Col-0 the same degree of resistance that was observed in wild-type *Ler-0* plants. It is unclear how transcripts derived from this clone would be polyadenylated and terminated. There is no obvious consensus polyadenylation signal within the intron 7 sequence. We presume that the absence of an exon 8 splice acceptor site causes termination of the *RPP5* protein at a TAA stop codon that lies at the start of the intron 7 sequence (Figure 3).

The *RPP5* Protein Is Most Similar to N and L6

RPP5 is most similar to the N and L6 proteins (see Figures 6A and 7) and less closely related to *RPS2* and *RPM1*. The high degree of sequence conservation between *RPP5*, N, and L6 extends to the coincident positions of the intron-exon splice junctions giving rise to exons 1, 2, and 3 of the full-length transcripts. This suggests that these particular *R* genes have evolved from a common ancestral gene (Pryor and Ellis, 1993; Ellis et al., 1995). In contrast to *RPP5*, N, and L6, both *RPS2* and *RPM1* lack introns and possess a predicted leucine zipper motif. Furthermore, the *RPS2* and *RPM1* proteins lack a region of similarity to the cytoplasmic domains of the *Drosophila* Toll and mammalian IL-1R receptor proteins (see discussion below), suggesting that they may interact with somewhat different signaling pathways.

RPP5, N, and L6 exhibit the highest level of sequence conservation within the first two exons but greater divergence in the C-terminal portion of their full-length proteins. In addition to the three consensus motifs (*kinase-1a* or P-loop, *kinase-2*, and *kinase-3a*) forming the NBS, exons 1 and 2 possess several other highly conserved regions (exon 1 is shown in Figure 7A). The functional significance of these is not known, although a G/SLPL sequence (*RPP5* amino acids 380 to 383 in Figure 3) is conserved among all *R* genes that have cytoplasmic LRRs, including *Prf* (Salmeron et al., 1996), and may therefore have an important functional role.

Whitham et al. (1994) noted previously that a region at the N terminus of the N protein (amino acids eight to 150 in both N and *RPP5*) resembles the cytoplasmic domains of the *Drosophila* Toll (Hashimoto et al., 1988) and mammalian IL-1R (Sims et al., 1989) transmembrane receptors. Our amino acid sequence comparisons and corresponding secondary structure predictions for *RPP5*, N, L6, Toll, and IL-1R (Figure 7) support their data, although in both analyses, IL-1R diverges from the other sequences at several points, suggest-

ing a degree of functional divergence. This is consistent with the conclusions of Hopp (1995), who proposed an additional GTPase function for the IL-1R cytoplasmic domain, based on key amino acid similarities with the ras GTPase superfamily. These conserved catalytic residues were not evident in Toll and are not present in the corresponding R protein domains.

IL-1R mediates a pathway leading to the induction of acute immune response genes in mammalian cells, whereas Toll controls the induction of genes that establish dorsoventral polarity in the *Drosophila* embryo. Recent studies established that a Toll-IL-1R pathway also controls the innate immune response of *Drosophila* adults to pathogen invasion (Ip et al., 1993; Hultmark, 1994; Rosetto et al., 1995), resulting in the transcriptional activation of antimicrobial peptide genes through a similar signaling process (Ip et al., 1993; Lemaitre et al., 1995, 1996). The striking degree of structural and functional conservation within these pathways suggests that they have coevolved from an ancient defense signaling mechanism (Fryxell, 1996). Therefore, it seems plausible that the N, L6, and *RPP5* proteins activate a signaling cascade related to the Toll and IL-1R pathways. Thus, the conserved N-terminal portion of these R proteins would confer the downstream "effector" capability, involving a Toll-like function as well as ATP or GTP binding.

It follows that the capability for disease resistance "specificity" is likely to reside in the more highly diverged C-terminal regions of *RPP5*, N, or L6, which contain the LRRs. Part of the LRR domain could direct specific binding of a pathogen-derived ligand. This is the anticipated role for the extracellular LRRs encoded by the tomato *Cf-9* (Jones et al., 1994) and *Cf-2* (Dixon et al., 1996) *R* genes, mediating resistance to fungal-derived Avr peptides that are present and active in the intercellular spaces of responding plants. Candidate *avr*-derived elicitors from *P. parasitica* have not yet been identified. At present, it is difficult to envisage how these might efficiently come in contact with a cytoplasmically localized R protein to initiate the defense pathway. The pathogen grows extracellularly but establishes intimate physical association with the plant plasma membrane (Koch and Slusarenko, 1990; Lucas et al., 1995), thus possibly allowing the passage of elicitor molecules into the plant cell. It is now known that the bacterial avirulence proteins AvrB (Gopalan et al., 1996) and AvrPto (Scofield et al., 1996; Tang et al., 1996) interact directly with their corresponding plant *R* gene products inside the plant cell, implicating active mobilization of the bacterial elicitors into the plant cell cytoplasm.

Molecular Evidence for the Evolution of New *RPP* Gene Specificities

Considerable genetic evidence exists for the clustering of many *R* genes in complex loci (Hulbert and Michelmore, 1985; Hulbert and Bennetzen, 1991; Jones et al., 1993;

Reignault et al., 1996; Holub and Beynon, 1997) and lends support to the hypothesis that the different *R* gene specificities might arise by unequal crossing-over or gene conversion events between related sequences (Hulbert and Bennetzen, 1991; Sudupak et al., 1993; Ellis et al., 1995; Richter et al., 1995). We show here the presence of a small highly polymorphic gene family at the *RPP5* locus (see Figure 1). Work is now in progress to characterize the remaining family members and examine the structure of this locus, spanning ~50 kb, in both the resistant and susceptible ecotypes, *Ler-0* and *Col-0*. This analysis will give important insights into the evolution of *RPP* genes.

Novel resistance specificities may also arise potentially through DNA rearrangements within the *R* genes themselves (Ellis et al., 1995; Lawrence et al., 1995; Jones and Jones, 1996). Examination of the repeated structure of the LRRs in *RPP5* provides molecular evidence for these evolutionary processes. Series of amino acids are repeated, as shown in Figure 4A. Four pairs of probable duplication products can be identified. Analysis of the corresponding nucleotide sequences confirms the occurrence of the inferred duplication events (data not shown). The coincident positions and the same sizes of introns 3 and 4 (see Figures 1C and 4) are also indicative of a recent duplication event. From the relationships between the repeats, we envisage that LRRs 5 to 10 most closely resemble a core ancestral LRR sequence from which the extended LRR domain was then derived. Interestingly, clearly reiterated sequences are not apparent in the *N* gene-encoded LRRs. *L6*, however, contains a tandem duplication of 140 amino acids toward the C terminus of its LRR domain (Lawrence et al., 1995). Duplications within the LRR domains of other plant proteins arising from probable intragenic recombination have been reported recently (Torii et al., 1996; Tornero et al., 1996), including the tomato *Cf-2* resistance gene product (Dixon et al., 1996). Together, these examples demonstrate how readily the LRR domain could evolve novel presentational surfaces for protein or peptide ligand binding (Kobe and Deisenhofer, 1994).

The *RPP5* mutant allele identified in FL-387a contains an intragenic duplication of four complete LRRs. This duplication may have arisen from an unequal crossing-over event between two sequences of identity in LRRs 13/14 and 17/18 (Figure 4B), either from the effects of the mutagen or spontaneously. Whatever the cause of the *RPP5* gene rearrangement, it again shows the capacity of the LRR domain to evolve novel configurations, and thus potentially new resistance specificities, in response to selection pressure from the pathogen. Similarly, two spontaneous flax mutations, X3A and X75, at the *L6* locus were shown to have small DNA insertions, although the nature of these rearrangements was not reported (Lawrence et al., 1995). The proposed expansion or contraction of the LRR domains within some *R* genes as a means of generating diversity is reminiscent of the variation in numbers of amino acid repeats determining plant host specificity in several bacterial *avr* genes (Herbers et al., 1992; Yang and Gabriel, 1995).

Expression of *RPP5* and Related Transcripts

The length of the *RPP5*-encoded mRNA (~4 kb) is consistent with the predicted full-length transcript and appeared to be of low abundance, at least in healthy *Ler-0* plants. Experiments are in progress to measure *RPP5* expression levels more precisely under unchallenged and pathogen-challenged conditions and at different developmental stages of the plant.

Alternative transcripts were previously identified for both the *N* and *L6* genes, which potentially give rise to the truncated gene products *N^{tr}* and *L6^{tr}* (Whitham et al., 1994; Dinesh-Kumar et al., 1995; Lawrence et al., 1995; see also Figure 6B). Both forms of truncated protein lack most of the LRR region and possess a small C-terminal extension beyond the amino acid sequence encoded by exons 1, 2, and 3. In this study, we have isolated cDNAs representing transcripts from one or more members of the *Ler-0* *RPP5* gene family that encode a truncated form of *RPP5* (*RPP5^{tr}*). *N^{tr}*, *L6^{tr}*, and *RPP5^{tr}* all contain the TIR (for *Toll*, *IL-1R*, *R* gene; Figure 6) and NBS domains but essentially lack the LRR domain. Three members of the *R* gene class encoding TIR-NBS-LRR products have now been cloned. For all three, transcripts encoding truncated proteins are produced, although by three different mechanisms (see Figures 6B and 6C). This suggests an important role for the truncated products, and it has been proposed that *N^{tr}* may function as a dominant regulator of the activity of the full-length *N* protein, possibly by stochastic binding of nucleotides (Dinesh-Kumar et al., 1995). It seems unlikely that the truncated proteins, lacking significant LRR domains, are involved in pathogen recognition. The latter conclusion is supported by the finding that *RPP5^{tr}* is encoded by a member of the *RPP5* gene family and not by the gene that determines specificity—*RPP5*.

We can now address how the *RPP5* gene product interacts with other plant signaling components. Mutational screens in *Arabidopsis* for plants altered in resistance to *Noco2* have revealed at least two other loci that are required for *RPP5* function (Parker et al., 1996; J.E. Parker, unpublished data). One of these, *EDS1*, is essential for downy mildew resistance mediated by several different *RPP* genes (Parker et al., 1996), suggesting a role in a convergent pathway, possibly downstream of specific *P. parasitica* recognition.

METHODS

Cultivation of *Arabidopsis thaliana* Plants and Pathogenicity Tests

The origins of Columbia (*Col-0*) and Landsberg *erecta* (*Ler-0*) are as reported by Parker et al. (1993). Ecotype Nossen (*No-0*) was a gift from F. Ausubel (Massachusetts General Hospital, Boston). *Ler-0* × *Col-0* F₂ seed were used for high-resolution mapping of *RPP5*. Fast neutron-bombarded *Ler-0* M₂ seed were obtained from Lehle Seeds (Round Rock, TX). The conditions for plant cultivation, maintenance

of *Peronospora parasitica* isolate Noco2, and pathogenicity tests were as described by Reignault et al. (1996).

Plant Genomic DNA Preparations

Large-scale plant genomic DNA preparations were made as described previously (Parker et al., 1993). Rapid, small-scale DNA samples were prepared as described by Reignault et al. (1996). The method of Klimyuk et al. (1993) was used to prepare denatured DNA suitable for polymerase chain reaction (PCR) amplification with specific oligonucleotide primers. The latter method was used to screen for recombinants among *Ler-0* × *Col-0* F₂ progeny. Genomic DNA from *Brassica oleracea* (Alboglabra), *B. napus* (109), and *B. campestris* (R-C-13) was a kind gift from M. Trick (John Innes Centre).

DNA Manipulations and Cosmid Markers

General methods for DNA manipulation were performed as described previously (Sambrook et al., 1989). Cosmid clones g13683 and g4539 (Schmidt and Dean, 1995) were cultured in the presence of 30 mg/L kanamycin. The *Col-0* genomic DNA clones 15D15 and 16N19, in the cosmid vector pLAFFR3 (C. Lister and C. Dean, unpublished data), were grown in the presence of 10 mg/L tetracycline. Cosmid DNA was prepared using the alkaline lysis method (Sambrook et al., 1989). Whole ³²P-labeled cosmid DNA was used to probe blots of plant genomic DNA or yeast colony filters.

Yeast Artificial Chromosome DNA Manipulations

³²P-labeled probes were hybridized to filters of four yeast artificial chromosome (YAC) libraries: EG and S (Grill and Somerville, 1991), EW (Ward and Gen, 1991), and YUP (Ecker, 1990), as detailed by Schmidt and Dean (1995). Intact yeast chromosomal DNA was prepared as described by Gibson and Somerville (1992). The DNA was separated by contour-clamped homogeneous field (CHEF) gel electrophoresis, using a CHEF-DR11 instrument (Bio-Rad), and transferred to a GeneScreen hybridization membrane (New England Nuclear Research Products, Boston, MA). Sizes of individual YAC clones were estimated by probing blots with ³²P-labeled pBR322 DNA. Left-end or right-end YAC DNA probes were generated by inverse PCR using YAC vector nested primers, as described previously (Gibson and Somerville, 1992; Schmidt and Dean, 1995). The inverse PCR-derived DNA products were agarose gel purified and then used to probe blots of digested plant genomic DNA or yeast chromosomal DNA. For the restriction enzyme mapping of YAC EW16B10, intact yeast chromosomal DNA carrying YAC EW16B10 was digested in agarose gel plugs with the following rare cutting restriction enzymes: BssHII, EagI, NarI, NruI, and SmaI. DNA fragments were then separated by CHEF gel electrophoresis and transferred to GeneScreen membranes, as described above. Hybridization with the EW16B10 left-end probe was used to confirm partial digestion and to establish the restriction map. Sequential hybridizations of the same blot were used to position additional markers.

PCR-Based Markers

C18a, a randomly amplified polymorphic DNA (RAPD) marker (Parker et al., 1993), was blunt-end cloned into pBluescript KS+. The de-

rived marker, C18, was used to probe blots of restriction enzyme-digested plant and yeast genomic DNA, YAC library filters, and a *Ler-0* genomic DNA cosmid library (see below). Codominant ecotype-specific PCR-based markers were generated for C18 and for the cosmid clone g4539. Double-stranded DNA sequencing was performed with M13 universal forward and reverse primers for the 570-bp C18 insert and for a polymorphic 800-bp HindIII fragment of g4539. The sequence information was used to design specific primers (information for these and SC5 is available at <http://genome-www.stanford.edu>; SC5 primers were kindly provided by M. Stammers [John Innes Centre]). Amplification conditions were essentially as described by Reignault et al. (1996).

Arabidopsis Transformation

C18 was also used to probe filters of an *Ler-0* genomic DNA library constructed in the binary cosmid vector pCLD04541 (Jones et al., 1992; C. Lister and C. Dean, unpublished data) and mobilized into *Agrobacterium tumefaciens* GV3101 (see Mindrinos et al., 1994). The method of Valvekens et al. (1988) for stable *Agrobacterium*-mediated root transformation was used to generate transformants of ecotype *No-0*. The whole plant infiltration method of Bechtold et al. (1993) was used for transformation of ecotype *Col-0*. Transformed (T₁) seedlings were selected on *Arabidopsis* medium (Bechtold et al., 1993) containing 50 mg/L kanamycin. Neomycin phosphotransferase II assays were performed with individual T₂ generation transformants, according to the method of McDonnell et al. (1987). Assessment of the segregation of kanamycin resistance and sensitivity in T₃ transformant progeny was performed on selective *Arabidopsis* medium (as above).

Nucleotide Sequence Determination and Analysis of pCLD29L17

A shotgun cloning approach was employed to determine the nucleotide sequence of cosmid pCLD29L17. *Escherichia coli* (DH5α) transformants harboring recombinant clones containing *Arabidopsis* DNA were identified by hybridization to pCLD29L17 insert DNA. M13 universal forward and reverse primers were employed to determine end sequences, using the Dye Deoxy terminator cycle sequencing method (Applied Biosystems [ABI], La Jolla, CA) and an ABI model 373A sequencing system. Sequence contigs were assembled using UNIX versions of the Staden programs package, including TED and XBAP (Roger Staden, MRC, Cambridge, UK). Specific primers were used to fill sequence gaps. Computer-aided sequence similarity searches were made with Blast (Altschul et al., 1990) programs and the National Center for Biotechnology Information (Bethesda, MD) nucleotide and peptide sequence databases. Analysis of primary sequences was performed using the University of Wisconsin Genetics Computer Group programs. Secondary structure predictions were made with the nnpredict program (at <http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html>).

Construction of a Subclone Containing *RPP5*

A subclone of pCLD29L17 was constructed in the binary vector SLJ7292 (nearly identical to SLJ44024; Jones et al., 1992). The subclone, designated pRPP5-2, contained a 6304-bp DNA fragment defined by a BglIII restriction enzyme site 1298 bp 5' to the putative

translation initiation codon and a PstI restriction enzyme site subsequently found to be 83 bp 5' to the translation termination codon. pCLD29L17 was mobilized into *Agrobacterium* and transferred to *Arabidopsis* Col-0, as described above.

mRNA and cDNA Analysis

For RNA gel blot analysis, poly(A)⁺ RNA was prepared from the aerial tissues of 20-day-old plants by using a Pharmacia Biotech mRNA purification kit. For each sample, 8 µg of poly(A)⁺ RNA was separated on a denaturing agarose gel (Sambrook et al., 1989) and transferred to Hybond-N (Amersham International, UK) membranes. Thirty- to 40-day-old *Ler-0* plants were taken to make two cDNA libraries from the poly(A)⁺ RNA (as above) by using a λZAPII cDNA synthesis kit (Stratagene). The second library involved size fractionation of the cDNA greater than ~1.8 kb. Plaque-forming units (10⁶) from each library were screened with a genomic probe extending from the NcoI site at the *RPP5* initiation codon to an XbaI site 1019 bp 5' to the presumed termination codon.

Reverse Transcriptase-PCR Analysis of Intron Splicing

First-strand cDNA was prepared from seedling leaf mRNA by using a cDNA Synthesis System Plus (Amersham). Amplification from 2.5 ng of cDNA was performed in a thermal cycler (model 9600; Perkin-Elmer Cetus, Norwalk, CT) for 40 cycles (94°C for 15 sec, 50°C for 30 sec, and 72°C for 1 min) in a standard PCR reagent mix containing 0.2 µM of each primer. Primer pairs flanking the predicted introns 1 to 7 were used in combination with primer pairs predicted to correspond to exon sequences. Cloned PCR products were sequenced as described above by using a combination of vector-specific and insert-specific primers. To detect incomplete processing of the *RPP5* transcript similar to that seen for *L6* (Lawrence et al., 1995), primer A1 (5'-AAGATCTTCTGAACGGGCC), containing 16 nucleotides complementary to the 5' flanking sequence of intron 2 and four nucleotides complementary to the 3' flanking sequence of intron 2, was used with primer A2 (5'-ATTCTTCAAGATTAATTCC) that primes from exon 4.

ACKNOWLEDGMENTS

We are grateful to Paul O'Hara for technical assistance and to Andrew Davis for photography. We also thank Patrick Bovill and Alan Cavill for operating the ABI373 automated sequencer. This work was supported by the Gatsby Foundation and by a European Community Training and Mobility of Researchers postdoctoral fellowship award to E.A.v.d.B. and grant No. BIOT CT 90-0207 to R.S. and C.D.

Received January 30, 1997; accepted April 14, 1997.

REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.
- Bechtold, N., Ellis, J., and Pelletier, G. (1993). *In planta Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C.R. Acad. Sci. Ser. III Sci. Vie* **316**, 1194–1199.
- Bell, C.J., and Ecker, J.R. (1994). Assignment of 30 microsatellite loci to the linkage map of *Arabidopsis*. *Genomics* **19**, 137–144.
- Bent, A.F., Kunkel, B.N., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J., and Staskawicz, B.J. (1994). *RPS2* of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. *Science* **265**, 1856–1860.
- Crute, I., Beynon, J., Dangl, J., Holub, E., Mauch-Mani, B., Slusarenko, A., Staskawicz, B., and Ausubel, F. (1994). Microbial pathogenesis of *Arabidopsis*. In *Arabidopsis*, E.M. Meyerowitz and C.R. Somerville, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 705–747.
- Dangl, J.L. (1993). Applications of *Arabidopsis thaliana* to outstanding issues in plant-pathogen interactions. *Int. Rev. Cytol.* **114**, 53–83.
- Dinesh-Kumar, S.P., Whitham, S., Choi, D., Hehl, R., Corr, C., and Baker, B. (1995). Transposon tagging of tobacco mosaic virus resistance gene *N*: Its possible role in the TMV-*N*-mediated signal transduction pathway. *Proc. Natl. Acad. Sci. USA* **92**, 4175–4180.
- Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.M., Harrison, K., and Jones, J.D.G. (1996). The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine rich repeat proteins. *Cell* **84**, 451–459.
- Ecker, J. (1990). PFGE and YAC analysis of the *Arabidopsis* genome. *Methods* **1**, 186–194.
- Ellis, J.G., Lawrence, G.J., Finnegan, E.J., and Anderson, P.A. (1995). Contrasting complexity of two rust resistance loci in flax. *Proc. Natl. Acad. Sci. USA* **92**, 4185–4188.
- Flor, H.H. (1971). Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* **9**, 275–296.
- Fryxell, K.J. (1996). The coevolution of gene family trees. *Trends Genet.* **12**, 364–369.
- Gibson, S.I., and Somerville, C. (1992). Chromosome walking in *Arabidopsis thaliana* using yeast artificial chromosomes. In *Methods in Arabidopsis Research*, C. Koncz, N.-H. Chua, and J. Schell, eds (Singapore: World Scientific), pp. 119–143.
- Gopalan, S., Bauer, D.W., Alfano, J.R., Loniello, A.O., He, S.Y., and Collmer, A. (1996). Expression of the *Pseudomonas syringae* avirulence protein AvrB in plant cells alleviates its dependence on the hypersensitive response and pathogenicity (Hrp) secretion system in eliciting genotype-specific hypersensitive cell death. *Plant Cell* **8**, 1095–1105.
- 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.
- Grill, E., and Somerville, C. (1991). Construction and characterization of a yeast artificial chromosome library of *Arabidopsis* which is suitable for chromosome walking. *Mol. Gen. Genet.* **226**, 484–490.
- Hammond-Kosack, K.E., and Jones, J.D.G. (1996). Resistance gene-dependent plant defense responses. *Plant Cell* **8**, 1773–1791.
- Hashimoto, C., Hudson, K.L., and Anderson, K.V. (1988). The *Toll* gene of *Drosophila*, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. *Cell* **52**, 269–279.

- Herbers, K., Conrads-Strauch, J., and Bonas, U.** (1992). Race-specificity of plant resistance to bacterial spot disease determined by repetitive motifs in a bacterial avirulence protein. *Nature* **356**, 172–174.
- Holub, E.B., and Beynon, J.L.** (1997). Symbiology of mouse-ear cress (*Arabidopsis thaliana*) and oomycetes. *Adv. Bot. Res.* **24**, 228–273.
- Holub, E.B., Beynon, J.L., and Crute, I.R.** (1994). Phenotypic and genotypic characterization of interactions between isolates of *Peronospora parasitica* and accessions of *Arabidopsis thaliana*. *Mol. Plant-Microbe Interact.* **7**, 223–239.
- Hopp, T.P.** (1995). Evidence from sequence information that the interleukin-1 receptor is a transmembrane GTPase. *Protein Sci.* **4**, 1851–1859.
- Hulbert, S.H., and Bennetzen, J.L.** (1991). Recombination at the *Rp1* locus of maize. *Mol. Gen. Genet.* **226**, 377–382.
- Hulbert, S.H., and Michelmore, R.W.** (1985). Linkage analysis of genes for resistance to downy mildew (*Bremia lactucae*) in lettuce (*Lactuca sativa*). *Theor. Appl. Genet.* **70**, 520–528.
- Hultmark, D.** (1994). Ancient relationships. *Nature* **367**, 116–117.
- Ip, Y.T., Reach, M., Engstrom, Y., Kadalayil, L., Cai, H., Gonzalez-Crespo, S., Tatei, K., and Levine, M.** (1993). *Dif*, a *dorsal*-related gene that mediates an immune response in *Drosophila*. *Cell* **75**, 753–763.
- Jones, D.A., and Jones, J.D.G.** (1996). The roles of leucine-rich repeat proteins in plant defenses. *Adv. Bot. Res.* **24**, 89–167.
- Jones, D.A., Dickenson, M.J., Balint-Kurti, P.J., Dixon, M.S., and Jones, J.D.G.** (1993). Two complex resistance loci revealed in tomato by classical and RFLP mapping of the *Cf-2*, *Cf-4*, *Cf-5*, and *Cf-9* genes for resistance to *Cladosporium fulvum*. *Mol. Plant-Microbe Interact.* **6**, 348–357.
- Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J., and Jones, J.D.G.** (1994). Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* **266**, 789–793.
- Jones, J.D.G., Shlumukov, L., Carland, F., English, J., Scofield, S., Bishop, G., and Harrison, K.** (1992). Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. *Transgenic Res.* **1**, 285–297.
- Klimyuk, V.I., Carroll, B.J., Thomas, C.M., and Jones, J.D.G.** (1993). Alkali treatment for rapid preparation of plant tissue for reliable PCR analysis. *Plant J.* **3**, 493–494.
- Kobe, B., and Deisenhofer, J.** (1994). The leucine-rich repeat: A versatile binding motif. *Trends Biochem. Sci.* **19**, 415–421.
- Koch, E., and Slusarenko, A.** (1990). *Arabidopsis* is susceptible to infection by a downy mildew fungus. *Plant Cell* **2**, 437–445.
- Konieczny, A., and Ausubel, F.M.** (1993). A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific PCR-based markers. *Plant J.* **4**, 403–410.
- Kunkel, B.N.** (1996). A useful weed put to work: Genetic analysis of disease resistance in *Arabidopsis thaliana*. *Trends Genet.* **12**, 63–69.
- Kyte, J., and Doolittle, R.F.** (1982). A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* **157**, 105–132.
- Lawrence, G.J., Finnegan, E.J., Ayliffe, M.A., and Ellis, J.G.** (1995). The *L6* gene for flax rust resistance is related to the *Arabidopsis* bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. *Plant Cell* **7**, 1195–1206.
- Lemaitre, B., Meister, M., Govind, S., Georgel, P., Steward, R., Reichhart, J.-M., and Hoffmann, J.A.** (1995). Functional analysis and regulation of nuclear import of *dorsal* during the immune response in *Drosophila*. *EMBO J.* **14**, 536–545.
- Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.-M., and Hoffmann, J.A.** (1996). The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell* **86**, 973–983.
- Lucas, J.A., Hayter, J.B.R., and Crute, I.R.** (1995). The downy mildews: Host specificity and pathogenesis. In *Pathogenesis and Host Specificity in Plant Diseases*, K. Kohmoto, U.S. Singh, and R.P. Singh, eds (Oxford, UK: Elsevier Science Ltd.), pp. 217–238.
- Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganai, M.W., Spivey, R., Wu, T., Earle, E.D., and Tanksley, S.D.** (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* **262**, 1432–1436.
- McDonnell, R.E., Clark, R.D., Smith, W.A., and Hinchee, M.A.** (1987). A simplified method for the detection of neomycin phosphotransferase II activity in transformed plant tissues. *Plant Mol. Biol. Rep.* **5**, 380–386.
- 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.
- Parker, J.E., Szabò, V., Staskawicz, B.J., Lister, C., Dean, C., Daniels, M.J., and Jones, J.D.G.** (1993). Phenotypic characterization and molecular mapping of the *Arabidopsis thaliana* locus, *RPP-5*, determining resistance to *Peronospora parasitica*. *Plant J.* **4**, 821–831.
- Parker, J.E., Holub, E.B., Frost, L.N., Falk, A., Gunn, N.D., and Daniels, M.J.** (1996). Characterization of *eds1*, a mutation in *Arabidopsis* suppressing resistance to *Peronospora parasitica* specified by several different *RPP* genes. *Plant Cell* **8**, 2033–2046.
- Pryor, T., and Ellis, J.** (1993). The genetic complexity of fungal resistance genes in plants. *Adv. Plant Pathol.* **10**, 281–305.
- Reignault, P., Frost, L.N., Richardson, H., Daniels, M.J., Jones, J.D.G., and Parker, J.E.** (1996). Four *Arabidopsis* *RPP* loci controlling resistance to the *Noco2* isolate of *Peronospora parasitica* map to regions known to contain other *RPP* recognition specificities. *Mol. Plant-Microbe Interact.* **9**, 464–473.
- Richter, T.E., Pryor, T.J., Bennetzen, J.L., and Hulbert, S.C.** (1995). New rust resistance specificities associated with recombination in the *Rp1* complex in maize. *Genetics* **141**, 373–381.
- Rosetto, M., Engström, Y., Baldari, C.T., Telford, J.L., and Hultmark, D.** (1995). Signals from the IL-1 receptor homolog, Toll, can activate an immune response in a *Drosophila* hemocyte cell line. *Biochem. Biophys. Res. Commun.* **209**, 111–116.
- Ryals, J., Uknes, S., and Ward, E.** (1994). Systemic acquired resistance. *Plant Physiol.* **104**, 1109–1112.
- Salmeron, J.M., Oldroyd, G.E.D., Rommens, C.M.T., Scofield, S.R., Kim, H.-S., Lavelle, D.T., Dahlbeck, D., and Staskawicz, B.J.** (1996). Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. *Cell* **86**, 123–133.

- Sambrook, J., Fritsch, E.F., and Maniatis, T.** (1989). *Molecular Cloning: A Laboratory Manual*, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
- Schmidt, R., and Dean, C.** (1995). Hybridization analysis of YAC clones. *Methods Mol. Cell. Biol.* **5**, 309–318.
- Scofield, S.R., Tobias, C.M., Rathjen, J.P., Chang, J.H., Lavelle, D.T., Michelmore, R.W., and Staskawicz, B.J.** (1996). Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. *Science* **274**, 2063–2065.
- Sherriff, C., and Lucas, J.A.** (1990). A study of the host range of isolates of *Peronospora parasitica* (Pers. ex. Fr.) from *Brassica* crop species. *Plant Pathol.* **39**, 77–91.
- Sims, J.E., Acres, R.B., Grubin, C.E., McMahan, C.J., Wignall, J.M., March, C.J., and Dower, S.K.** (1989). Cloning the interleukin 1 receptor from human T cells. *Proc. Natl. Acad. Sci. USA* **86**, 8946–8950.
- Song, W.-Y., Wang, G.-L., Chen, L.-L., Kim, H.-S., Pi, L.-Y., Holsten, T., Gardner, J., Wang, B., Zhai, W.-X., Fauquet, C., and Ronald, P.** (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* **270**, 1804–1806.
- Staskawicz, B.J., Ausubel, F.M., Baker, B.J., Ellis, J.G., and Jones, J.D.G.** (1995). Molecular genetics of plant disease resistance. *Science* **268**, 661–667.
- Sudupak, M.A., Bennetzen, J.L., and Hulbert, S.H.** (1993). Unequal exchange and meiotic instability of disease-resistance genes in the *Rpl* region of maize. *Genetics* **133**, 119–125.
- Tang, X., Frederick, R.D., Zhou, J., Halterman, D.A., Jia, Y., and Martin, G.B.** (1996). Initiation of plant disease resistance by physical interaction of *AvrPto* and *Pto* kinase. *Science* **274**, 2060–2063.
- Tör, M., Holub, E.H., Brose, E., Musker, R., Gunn, N., Can, C., Crute, I.R., and Beynon, J.L.** (1994). Map positions of three loci in *Arabidopsis thaliana* associated with isolate-specific recognition of *Peronospora parasitica* (downy mildew). *Mol. Plant-Microbe Interact.* **7**, 214–222.
- Torii, K.U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R.F., and Komeda, Y.** (1996). The *Arabidopsis* *ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell* **8**, 735–746.
- Tornero, P., Mayda, E., Gomez, M.D., Canas, L., Conejero, V., and Vera, P.** (1996). Characterization of LRP, a leucine-rich repeat (LRP) protein from tomato plants that is processed during pathogenesis. *Plant J.* **10**, 315–330.
- Traut, T.W.** (1994). The functions and consensus motifs of nine types of peptide segments that form different types of nucleotide-binding sites. *Eur. J. Biochem.* **229**, 9–19.
- Valvekens, D., Van Montagu, M., and Van Lijsebettens, M.** (1988). *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana* root explants by using kanamycin selection. *Proc. Natl. Acad. Sci. USA* **85**, 5536–5540.
- Ward, E.R., and Gen, G.C.** (1990). Isolation of single-copy-sequence clones from a yeast artificial chromosome library of randomly sheared *Arabidopsis thaliana* DNA. *Plant Mol. Biol.* **14**, 561–568.
- Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B.** (1994). The product of the tobacco mosaic virus resistance gene *N*: Similarity to Toll and the interleukin-1 receptor. *Cell* **78**, 1011–1115.
- Yang, Y., and Gabriel, D.W.** (1995). Intragenic recombination of a single plant pathogen gene provides a mechanism for the evolution of new host specificities. *J. Bacteriol.* **177**, 4963–4968.
- Zhou, J., Loh, Y.-T., Bressan, R.A., and Martin, G.B.** (1995). The tomato gene *Pti1* encodes a serine/threonine kinase that is phosphorylated by *Pto* and is involved in the hypersensitive response. *Cell* **83**, 925–935.