Theor Appl Genet (2002) 104:592–600

M.F. Bouzidi · S. Badaoui · F. Cambon · F. Vear
D. Tourvielle De Labrouhe · P. Nicolas · S. Mouzeyar

Molecular analysis of a major locus for resistance to downy mildew in sunflower with specific PCR-based markers

Received: 9 April 2001 / Accepted: 10 August 2001

Abstract Resistance of sunflower to the obligate parasite *Plasmopara halstedii* is conferred by specific dominant genes, denoted *Pl*. The *Pl6* locus confers resistance to all races of *P. halstedii* except one, and must contain at least 11 tightly linked genes each giving resistance to different downy mildew races. Specific primers were designed and used to amplify 13 markers covering a genetic distance of about 3 cM centred on the *Pl6* locus. Cloning and sequence analysis of these 13 markers indicate that *Pl6* contains conserved genes belonging to the TIR-NBS-LRR class of plant resistance genes.

Keywords Disease resistance · *Helianthus annuus* · *Plasmopara halstedii* · Molecular markers · Toll-Interleukin-1 Receptor (TIR) · Nucleotide binding site (NBS)

Introduction

Downy mildew caused by *Plasmopara halstedii* is one of the most important diseases of cultivated sunflower (*Helianthus annuus* L). Genetic studies of resistance to this parasite have shown that major dominant genes, denoted *Pl*, control resistance to different races of *P. halstedii*. So far, ten *Pl* genes have been described; Miller and Gulya (1991) made crosses with wild *Helianthus* and described three *Pl* genes denoted *Pl6*, *Pl7* and *Pl8*. *Pl6* was obtained from wild *H. annuus*, whereas *Pl7* came from *Helianthus praecox* and *Pl8* from *Helianthus argophyllus*. One of the most striking features of these genes was that each one of them conferred resistance to all races of *P. halstedii* known at that time, suggesting that either they conferred some non-race-specific, complete resistance, or that these "genes" were in fact complex loci containing several linked *Pl* genes giving resistance to individual races.

Artificial infection of 150 F3 progenies of a cross between a sunflower line containing *Pl6* and a susceptible line, by five different *P. halstedii* races, showed that the *Pl6* locus could be split into at least two genetically distinct regions, one giving resistance to races 100 and 300 and a second giving resistance to races 700, 703 and 710 (Vear et al. 1997). This was the first report of the clustering of *P. halstedii* resistance genes in sunflower. The *Pl6* locus may however contain at least 11 functional *Pl* genes since the sunflower lines which contain this locus are resistant to 11 races of downy mildew.

In recent years, many different disease-resistance genes (R-genes) have been cloned from plants and they confer resistance to fungal, viral, bacterial, insect and nematode pathogens. These genes contain conserved domains that can account for many of the predicted functions of R-genes. Five classes of R-genes are now recognised (Ellis and Jones 1998; Martin 1999; Meyers et al. 1999): intracellular protein kinases; receptor-like protein kinases with an extracellular leucine-rich repeat (LRR) domain; intracellular LRR proteins with a nucleotide binding site (NBS) and a leucine zipper (LZ) motif; intracellular NBS-LRR proteins with a region with similarity to the Toll and interleukin-1 receptor (TIR) proteins; and LRR proteins that code for membrane-bound extracellular proteins. Despite these significant insights into R-gene structure, much remains to be elucidated about the molecular mechanisms by which R-proteins recognise pathogens and transduce this information in the plant cell to initiate defence responses.

In sunflower, Gentzbittel et al. (1998) used degenerate primers designed from the conserved nucleotide binding domains of *N* from tobacco (Whitham et al.

Communicated by M.A. Saghai Maroof

M.F. Bouzidi · S. Badaoui · F. Cambon · P. Nicolas
S. Mouzeyar (✉)
UMR 1095 INRA-UBP "Amélioration et santé des plantes", Université Blaise Pascal.
24, Avenue des Landais 63177 Aubière Cedex, France
e-mail: Said.MOUZEYAR@ovgv.univ-bpclermont.fr
Fax: (+33)-4-73-40-79-14

F. Vear · D.T. de Labrouhe
INRA 234, avenue du Brézet 63039 Clermont Ferrand Cedex 02, France

1994), *RPS2* from *Arabidopsis thaliana* (Mindrinos et al. 1994) and *L6* from flax (Lawrence et al. 1995). The resulting amplification products were shown to be members of a multigene family. One clone was sequenced and mapped close to the *Pl6* cluster for resistance to downy mildew. Sequence analysis of this resistance-gene analog (RGA) showed considerable homology with the nucleotide binding domains of previously cloned resistance genes in other species.

In this paper, we report further analysis of the complexity of the *Pl6* locus. In order to obtain more molecular markers suitable for positional cloning of *Pl* genes, specific primers designed for the N-terminal region of a sunflower RGA sequence were used. Thirteen specific markers [sequence-tagged sites (STSs)] for the *Pl6* locus were cloned, sequenced, and mapped to a genetic distance of 3 cM. Evolutionary relationships between these sequences are discussed.

Materials and methods

Plant material and disease evaluation

The USDA sunflower line HA335 contains the *Pl6* locus (Miller and Gulya 1991) and is resistant to all major races of *P. halstedii*. The sunflower line H52 (from ARS, South Africa) is susceptible to all known races of *P. halstedii*. These lines were crossed and a segregating population of 142 F2 individuals was obtained. The corresponding F3 families were then tested for resistance to five races of *P. halstedii*, making it possible to classify the F2 plants as homozygous resistant, homozygous susceptible or heterozygous (Vear et al. 1997). Young leaf tissue from the F_2 plants was collected and DNA was isolated by freeze-drying followed by CTAB extraction, as described by Saghai Maroof et al. (1984). DNA from 12 homozygous susceptible or resistant lines of the F_2 population were pooled to form bulk-susceptible and bulk-resistant samples.

Development of PCR-based specific markers or STSs

In order to detect and develop PCR-markers, we chose to proceed by two stages: initially we obtained cDNA sequences by RACE-PCR; in a second stage, we used these sequences as templates for designing specific primers for cloning genomic sequences. This methodology is summarised in Fig. 1.

5′ and 3′ rapid amplification of cDNA ends (RACE) of RGA in sunflower

RNA was extracted from sunflower line HA335 using the method described by Bogorad et al. (1983), and the polyA with the PolyA-Tract mRNA Isolation System (Promega). Based on the sequence of the 627 bp-length RGA product (HA-NBSR3, accession number U96642) obtained with degenerate primers (Gentzbittel et al. 1998), 5′ and 3′ ends of the cDNA were obtained using the "Marathon cDNA amplification kit" (Clontech, Ozyme France). The amplified products were cloned into pGEM-T Easy vector (Promega) and sequenced by Genome Express (Grenoble, France).

Amplification of full-length genomic RGA

To clone genomic fragments with homology to plant resistance genes, we used specific primers from 5′(5′GGTAATGGCTGTTGAATTTATGGAGC3′, containing ATG, the initiation codon) and 3′ (5′TGTTGCCCATGGACCATTGATCC3′) portions of one 5′ and one 3′ RACE-PCR product, and amplified several genomic sequences. The PCR amplifications were carried out with 50 ng of sunflower DNA in the presence of 0.2 mM of each dNTP, 1 U of *Taq* DNA polymerase (Advantage 2, Clontech), 1×*Taq* polymerase buffer and 0.5 µM of each primer. PCR was carried out in a 2400 Perkin-Elmer thermocycler under the following conditions: 35 cycles for 10 s at 94°C, 30 s at 58°C (primer annealing), and 1 min 30 s at 72°C (primer extension). PCR products were separated using standard TAE agarose gel electrophoresis.

Amplification and cloning of polymorphic PCR fragments (STSs)

Primer pairs were designed (see Fig. 1) and tested for polymorphism using the bulked segregant analysis method as described by Michelmore et al. (1991). Three specific primer pairs were selected using sunflower RGAs, obtained as described in Table 1. These primers are all located at the 5′ end of these RGAs. The DNA fragments were amplified by polymerase chain reaction (PCR), using the 22–27-mer primers at a concentration of 0.5 µM. Amplification was performed as described above. PCR was carried out under the following conditions: 33 cycles for 10 s at 94°C, 30 s at 60°C, and 1 min 30 s at 72°C. Each amplified fragment was individually excised and purified from the gel using the GFX PCR purification system (Amersham-Pharmacia-Biotech, France). The purified fragments were cloned and sequenced.

Sequencing and sequence analysis

For each STS and for the full-length genomic fragment, one clone was chosen randomly and sequenced using the Dye-Terminator method (Genome Express, France). Both strands were sequenced and gaps were filled by primer walking. Each region of these clones was sequenced between two and five times. Sequences were subjected to data bank analysis using the BLAST algorithms (Altschul et al. 1997). Identification of the domain structure of the proteins was performed using the Pfam database (Bateman et al. 2000). Protein sequences were predicted using the Genscan program (Burge and Karlin 1997) and hand checked, then aligned with CLUSTALX computer software (Thompson et al. 1997). Alignments were shaded using the Genedoc software (Nicholas et al. 1997). A distance tree was obtained with CLUSTALX using the neighbor-joining (NJ) method. The CLUSTALX default options were used in the initial alignment.

Reverse transcriptase-PCR (RT-PCR) analysis of the expression of the STSs

To check whether the cloned STSs are parts of genes with potential transcriptional activity, accumulation of the transcripts was monitored using RT-PCR. The RNA was extracted both from H52 and HA335 healthy 12-day old hypocotyls. Five micrograms of total RNA were subjected to 3 units of DnaseI in the presence of Rnasine. Five microliters of Dnase-treated RNA were primed with an 18-mer oligodT and the first cDNA strand was synthesized using 200 units of reverse transcriptase (Superscript RT, Gibco-BRL, France). Because the level of the expression of these STSs was not known, several dilutions of the cDNA were used as templates in PCR, together with the three primer pairs described above, using the same PCR conditions as for the amplification of the STSs. RT-PCR products were separated using standard TAE agarose-gel electrophoresis.

Mapping the STSs

Marker orders and genetic distances were calculated with MAPMAKER 3.0 b software (Lander et al. 1987), to construct a map of the *Pl6* region, using a LOD value threshold of 3.0 and a maximum recombination fraction of 50. Each STS was scored as a

dominant marker (either present or absent) and the 13 STSs were mapped using 142 F2 individuals segregating for resistance to five races of downy mildew (Vear et al. 1997). The closest polymorphic RFLP marker, S124E1–2, was used to assign the STS to linkage group 1. Since the disease evaluations were made on F_3 families, the resistance genes were scored as co-dominant markers. A subset of markers consisting of the co-dominant markers (S124E1–2 and the downy mildew resistance genes) and two dominant markers from each parent was used to construct a core map of the *Pl6* region. The TRY command was employed to map the remaining markers. The best order was confirmed using the RIPPLE command of the software.

Results

Development of the STSs

Cloning of the 5′ and 3′ ends of RGA

The sequence of one partial RGA product obtained with degenerate primers (Gentzbittel et al. 1998) was used as a template to isolate the remaining 5′ and 3′ portions of the gene. The two longest clones were selected and the single-pass sequenced. Subsequent analysis and comparison with other plant resistance genes, such as the *N* gene of tobacco (Whitham et al. 1994), suggest that the 5′ clone (1,162-bp long, sequence not shown) contains one putative initiation codon ATG at position 15 and a 5′UTR consisting of 14 nucleotides. No attempt was made to obtain other RACE-PCR products containing larger 5′UTR sequences. The second clone, or the 3′ clone (1,839-bp long, sequence not shown) corresponding to the 3′ region, contains one putative TAA stop codon and a polyA tail. The two clones overlap over 61 bases and entirely cover the partial genomic clone obtained by Gentzbittel et al. (1998). However, the overlapping segment contains four mismatches indicating that the two clones may be part of different genes. In addition, numerous mismatches with the clone HA-NBSR3 (Gentzbittel et al. 1998) were detected, which may be due to the fact that this clone was obtained from RHA266, a sunflower line different from those used in this study.

Amplification of full-length genomic RGA

Using PCR and specific primers from the 5′ and 3′ portions of the RACE-PCR products, we amplified several genomic sequences (Fig. 1). The primers were homologous to the putative initiation and stop codons, and amplified bands ranging in size from 3.5 to 4.1 kb. The longest one (4,187 bp, accession number AF316405) was entirely sequenced. This RGA contains five exons and four introns (within the TIR domain, prior to the NBS domain, prior to the LRR and within the LRR domain). The predicted protein structure of this RGA genomic clone, determined using the Pfam database, is 770 amino-acids long and shares complete signature domains with some plant R-genes: TIR, NBS and LRR domains. The combination of these domains places this sunflower RGA in the tobacco *N* (Whitham et al. 1994), flax *L6* (Lawrence et al. 1995), and *A. thaliana RPP1* (Botella et al. 1998) family of R-genes.

Amplification and cloning of polymorphic PCR fragments (STSs)

Only three pairs of specific primers (Table 1) were tested and found to amplify polymorphic fragments between the susceptible sunflower line H52 and the resistant line HA335. Bulked segregant analysis (Michelmore et al. 1991) was used to detect markers potentially linked to

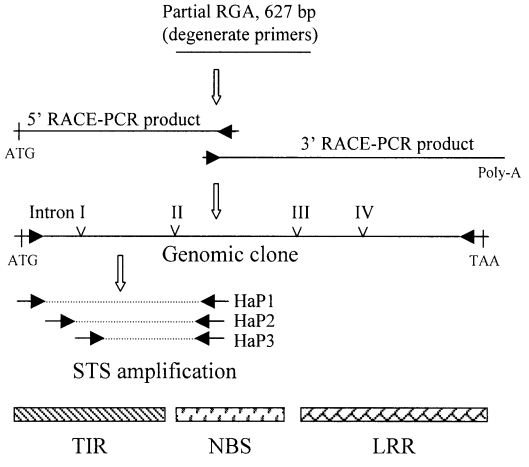

Fig. 1 Methodology used to develop STS markers. 5′ and 3′ rapid amplification of cDNA ends (RACE) of partial RGA (627-bp, accession number U96642) were obtained. These sequences were used as templates for designing specific primers indicated with *arrows* to amplify the genomic clones and STS markers

Table 1 Sequence of forward and reverse primers used to amplify 13 STSs linked to the *Pl6* locus. Their locations within the sunflower RGA accession number AF316405 are also indicated

Primer pair	Forward primer sequences	Reverse primer sequences
HaP1	5′GGTAATGGCTGTTGAATTTATGGAGC3′ Nucleotides 1 to 27	5′AGCATGATCCGGCTAGAGCCTTCTA3′ Nucleotides 2,162 to 2,137
HaP2	5′GTCTACTACATGGTTTCCGTTTTC3′ Nucleotides 104 to 127	5′TGCTTCTTCCTTCTATCTCACTC3′ Nucleotides 2,067 to 2,045
HaP3	5′GTTTGTGGATCATCTCTATGCG3′ Nucleotides 693 to 714	5′TGCTTCTTCCTTCTATCTCACTC3′ Nucleotides 2,067 to 2,045

five *P. halstedii* resistance genes (Fig. 2). Amplification with the primer pair HaP1 gave two major polymorphic bands between the two parents and the bulks (Fig. 2A); these bands were purified, cloned and sequenced; two additional bands were detected, one which was not polymorphic (about 2,100-bp long) was not cloned, and a faint band (1,610-bp large and denoted Ha-NBS15) was cloned and sequenced and shows homology with the NL25 sequence from potato (Hehl et al. 1999) (BLASTX p-value 6e-40). However, because it was faint, this band was not mapped. The primer pair HaP2 gave eight polymorphic bands between the two parents and the bulks (Fig. 2B); all the bands except the largest one (about 2,500 bp in the resistant parent) were purified, cloned and sequenced. The primer pair HaP3 gave four polymorphic bands between the two parents and the bulks (Fig. 2C); each amplified fragment was purified, cloned and sequenced. The size of each fragment is listed in Table 2.

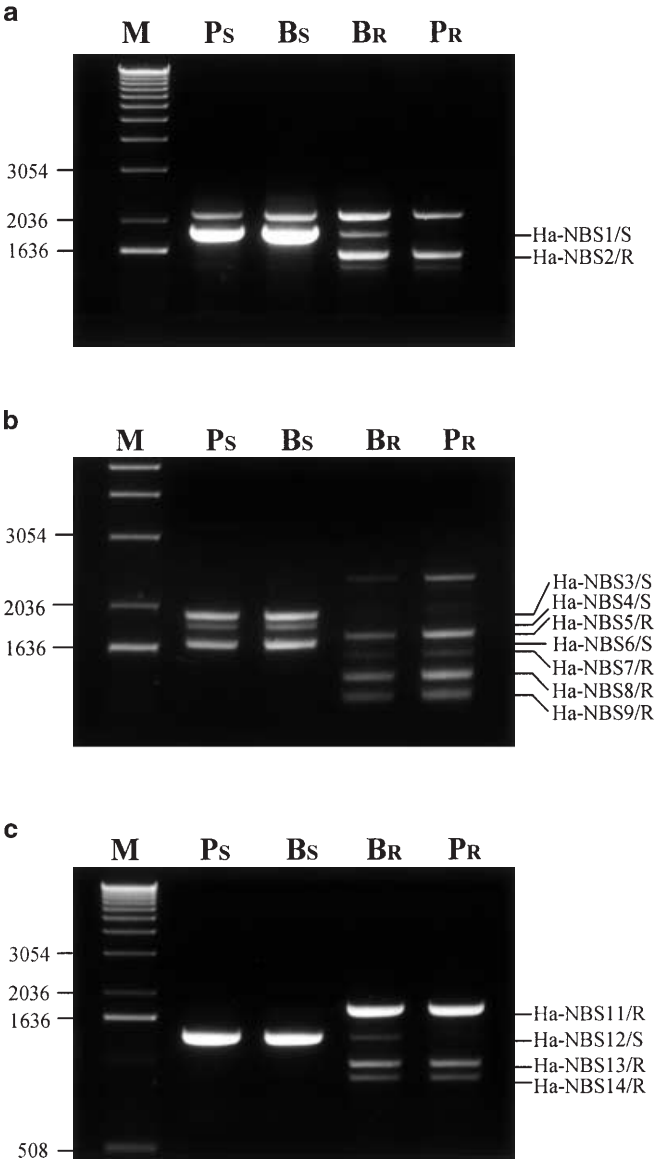

Fig. 2A–C STS amplification patterns. P_S susceptible parent (H52 line), B_S susceptible bulk; B_R resistant bulk and P_R resistant parent (HA335 line). *Lane M* 1-kb DNA ladder (Life Technologies). Primer combinations are as follows: **A** STS amplification patterns with HaP1 primers, **B** STS amplification patterns with HaP2 primers and **C** STS amplification patterns with HaP3 primers. The polymorphic fragments are listed on the right. *Ha-NBS/S* the sequence originating from the susceptible parent; *Ha-NBS/R* the sequence originating from the resistant parent. The bulk B_R included one F2 plant recombined for the marker Ha-NBS1/S and one for Ha-NBS12/S

Sequence analysis

Homology searches in gene databases using the BLAST suite (Altschul et al. 1997) indicate that the 13 STSs are all homologous to the TIR-NBS region of the TIR-NBS-LRR class of plant R-genes. The sequences showing the lowest *p*-value (i.e. the highest scores) when the BLASTX program was used, were NL25 (accession number, AJ009719) (*p*-value range: 1e-33 to 9e-43) and NL27 (accession number, AJ009720) (*p*-value range: 2e-33 to 5e-44) clones from potato (Hehl et al. 1999), the *N* gene (accession number, U15605) (*p*-value range: 8e-33 to 1e-36) for resistance to TMV in tobacco (Whitham et al. 1994), the *L6* gene (accession number, U27081) (*p*-value range: 1e-31 to 4e-35) for resistance to flax rust (Lawrence et al. 1995) and the *RPP1* gene (accession number, AF098962) (*p*-value range: 2e-29 to 1e-34) for resistance to downy mildew in *A. thaliana* (Botella et al. 1998). Analysis of the complete genomic RGA revealed that this sequence is homologous with the tobacco resistance gene *N* with a BLASTX *p*-value of 5e-64. The 5′ RACE clone shares homology with the TIR-NBS part of the NL27 cDNA clone from potato with a BLASTX *p*-value 5 e-51. The 3′ RACE clone shares homology with the C-terminal region of the *N* gene from tobacco with a BLASTX *p*-value of 5e-51.

Comparison of the 13 STSs

Due to the position of the primers used in this study, it was expected that amplification using different primer pairs would give different product sizes. The 13 STSs were trimmed so that only the most-internal and common parts corresponding to the position of the HaP3 primer pair were retained for sequence comparisons and alignments. Alignment of deduced amino acids of the internal and common sequence of the 13 STSs show that these sequences are conserved but not identical. They were compared with sunflower RGA (full-length genomic RGA), two R-genes (*N* and *L6*) and one RGA from potato (NL27). The similarities were especially close with one NBS motif (kinase 1a or the P-loop) and the conserved domains TIR-2 and TIR-3 within the N-terminal region (Fig. 3) (Meyers et al. 1999). This overall similarity, and the existence of the P-loop, TIR-2 and TIR-3 motifs, indicate that the STSs and the full-length genomic sunflower RGA belong to the TIR-NBS-LRR

Table 2 Origins and sizes of 13 STS markers within the *Pl6* locus

STS codes	Primer pairs	Parent	PCR product sizes	Accession numbers
Ha-NBS 1	HaP1	Susceptible	1,901 bp	AF316406
Ha-NBS 2	HaP1	Resistant	1,484 bp	AF316407
Ha-NBS 3	HaP2	Susceptible	1,694 bp	AF316408
Ha-NBS 4	HaP2	Susceptible	1,979 bp	AF316409
Ha-NBS 5	HaP2	Resistant	1,763 bp	AF316410
Ha-NBS 6	HaP2	Susceptible	1,700 bp	AF316411
Ha-NBS 7	HaP2	Resistant	1,589 bp	AF316412
Ha-NBS 8	HaP2	Resistant	1,414 bp	AF316413
Ha-NBS 9	HaP2	Resistant	1,260 bp	AF316414
Ha-NBS 11	HaP3	Resistant	1,811 bp	AF316415
Ha-NBS 12	HaP3	Susceptible	1,406 bp	AF316416
Ha-NBS 13	HaP3	Resistant	1,119 bp	AF316417
Ha-NBS 14	HaP3	Resistant	988 bp	AF316418

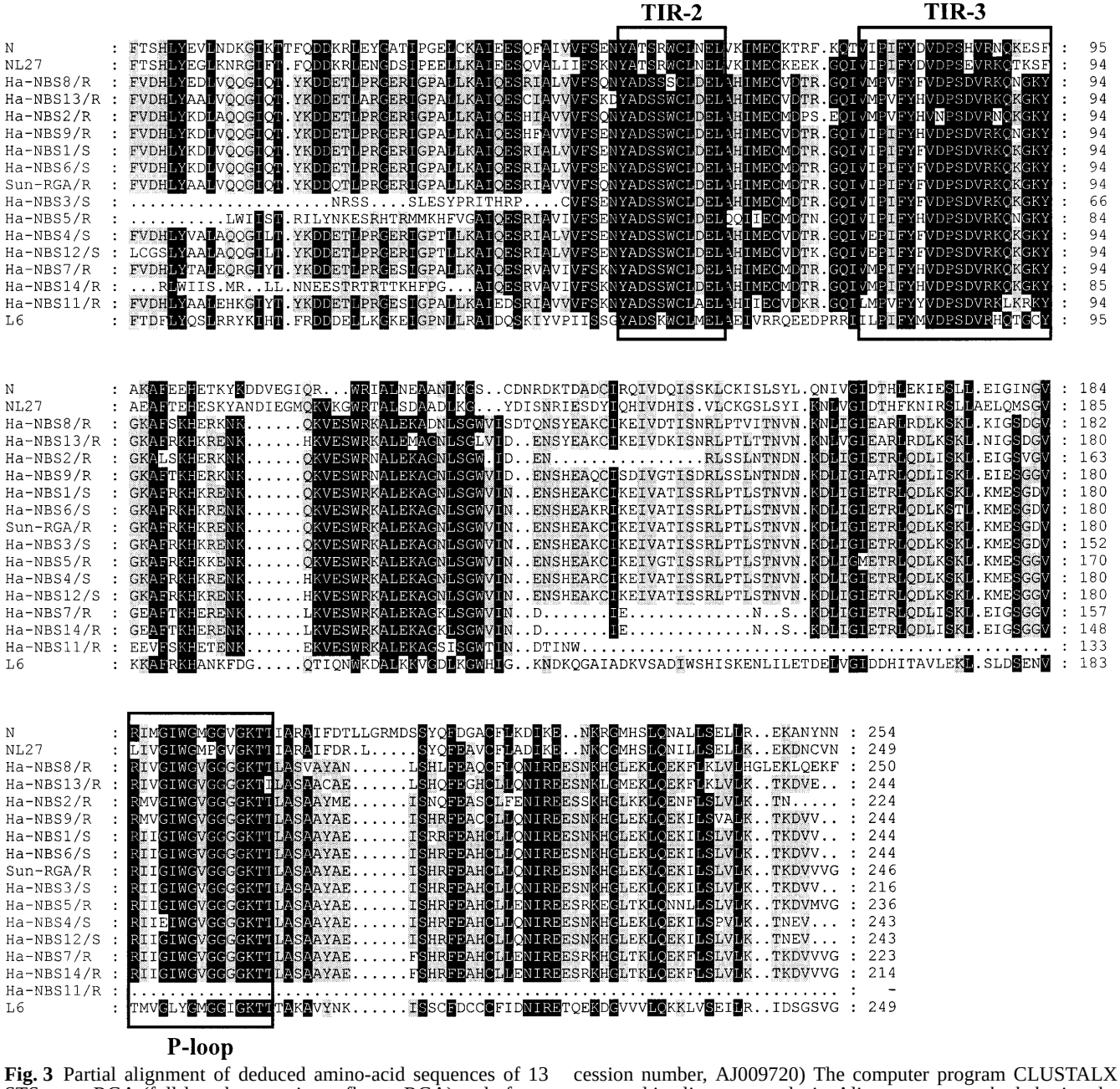

Fig. 3 Partial alignment of deduced amino-acid sequences of 13 STSs, sun-RGA (full-length genomic sunflower RGA) and of two R-genes, tobacco *N* (accession number, U15605) and flax *L6* (accession number, U27081), and one clone (NL27) from potato (accession number, AJ009720) The computer program CLUSTALX was used in alignment analysis. Alignments were *shaded* using the Genedoc software. The kinase1a (P-loop), TIR-2 and TIR-3 domains are *boxed*

Table 3 Percent amino-acid sequence identities of 13 STSs compared with sunflower RGA (Sun) and three R-genes (*N*, *NL27* and *L6*). Values were calculated using the CLUSTALX program with a gap opening penalty=10 and a gap extension penalty=0.05

	N	*NL27*	8/R	13/R	2/R	9/R	1/S	6/S	Sun/R	3/S	5/R	4/S	12/S	7/R	14/R	11/R
NL27	65															
8/R	43	44														
13/R	42	44	82													
2/R	39	40	70	71												
9/R	44	46	78	79	79											
1/S	43	44	81	83	75	87										
6/S	42	44	81	82	74	86	98									
Sun/R	42	44	81	83	73	85	96	96								
3/S	35	37	64	67	57	70	81	80	80							
5/R	37	39	64	67	60	70	78	77	79	77						
4/S	42	43	79	81	73	84	93	93	92	77	75					
12/S	41	43	78	81	72	82	92	92	91	78	76	96				
7/R	39	44	69	71	75	76	76	76	78	62	65	75	73			
14/R	33	38	59	61	65	65	66	66	68	64	66	65	65	86		
11/R	22	24	38	40	40	38	39	39	40	25	28	40	38	46	33	
L6	33	36	37	37	39	41	40	39	41	32	34	39	38	38	32	22

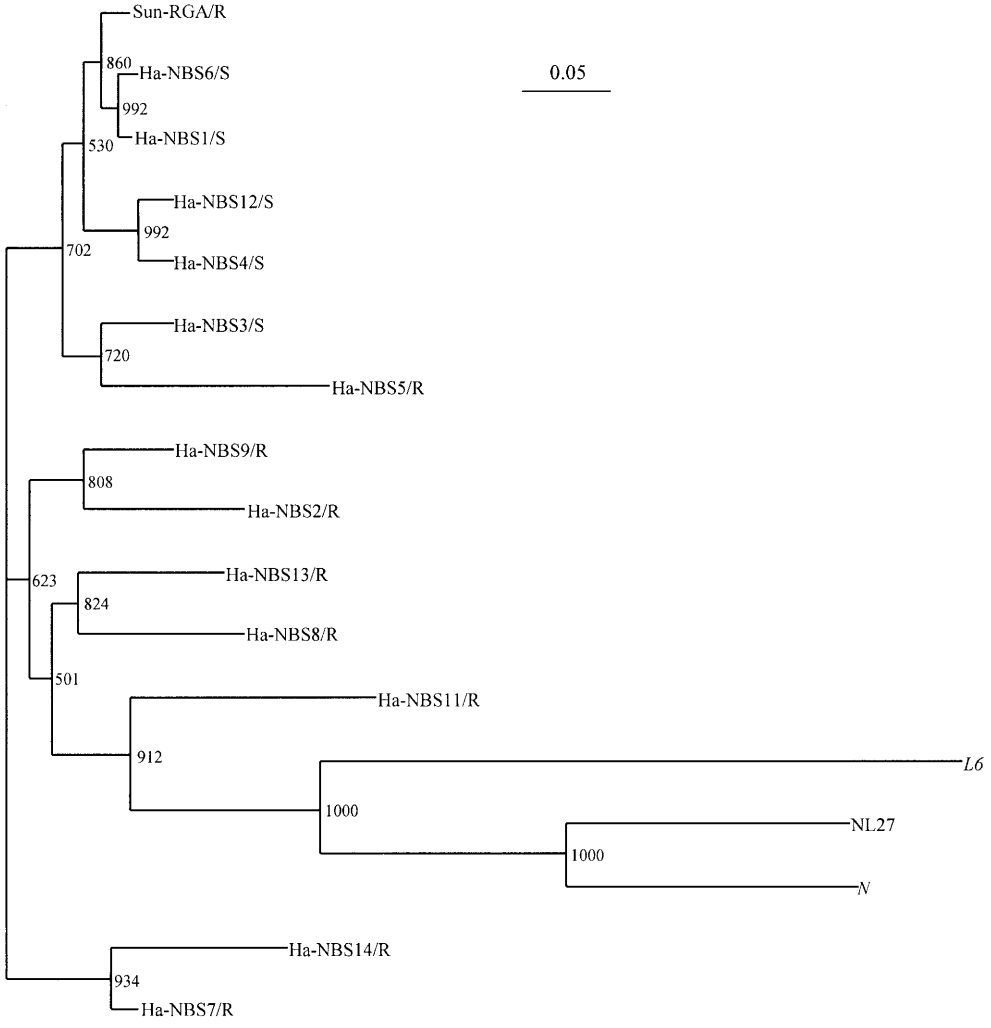

Fig. 4 Phylogenetic tree based on partial alignment of deduced amino-acid sequences of 13 STS markers, sunflower RGA and of two R-genes (*N* and *L6*) and one RGA (NL27). The tree was constructed using the neighbor-joining method provided in CLUSTALX. Ha-NBS/S stands for sequences originating from the susceptible parent H52; Ha-NBS/R stands for sequences originating from the resistant parent HA335

resistance-gene superfamily. The percent amino-acid sequence identities of these 13 STSs compared with previously cloned plant R-genes, or R-gene analogs, are listed in Table 3. The identities between the 13 STSs and the full-length sunflower RGA ranged from 57 to 98%, except for Ha-NBS11 which showed a percent identity of less than 46%. Twelve STSs and the full-length sunflower RGA were also similar to the *N* gene of tobacco, the NL27 sequence of potato and the *L6* gene of flax (average identity was 40.1%, 42.3% and 37.4% respectively).

In contrast, Ha-NBS11 showed less similarity with the *N* gene (22%), the NL27 clone (24%) and the *L6* gene (22%). Phylogenetic analysis was also performed to evaluate further relationships between sunflower RGAs and plant R-genes. The deduced amino-acid sequences of the 13 STSs, Sun-RGA and three plant R-genes were aligned and a neighbor-joining tree was generated from the alignment. Several iterations were performed and gave trees similar to that in Fig. 4. The majority of the nodes were found in at least 70% of 1,000 replicates in bootstrap analysis. Three major clusters were detected. The first contained all the STSs cloned from the susceptible parent, together with Ha-NBS5/R and the full-length RGA (Sun-RGA/R) from the resistant parent. The second cluster contained several STSs from the resistant parent and the resistance genes *L6* of flax, *N* of tobacco and NL27 of potato. The third cluster contained only two STSs, Ha-NBS7/R and Ha-NBS14/R, from the resistant parent.

RT-PCR analysis of the expression of the STSs

Among the three primer pairs tested, the HaP3 primer pair gave a faint band at approximately 700 bp when the cDNA from Ha335 was used as a template. The specificity of the amplification was confirmed by Southern blotting and hybridization with the 5′-RACE product. No additional bands were observed, even when the priming temperature was decreased to 55°C and the cycling number increased up to 40. The other primer pairs (HaP1 and HaP2) failed to amplify any detectable band even when the PCR products were subjected to Southern hybridization with the STSs as radioactively marked probes (data not shown).

Mapping of the STS markers

For the linkage group assignment of the STSs, the use in this study of the co-dominant RFLP marker S124E1–2, the closest polymorphic RFLP marker available (Gentzbittel et al. 1999), clearly demonstrated that all the 13 STSs map to the distal region of linkage group 1 of the RFLP composite map developed by Gentzbittel et al. (1999). Subsequent mapping of the STSs showed that they are tightly linked and lie within the *Pl6* locus containing the genes giving resistance to races 100, 300, 700, 703 and 710 of *P. halstedii*. Thus the *Pl6* locus may contain several copies of R-genes of the TIR-NBS-LRR class. The STS markers are clustered within a genetic distance of about 3 cM. Markers originating from the susceptible parent form two groups close to the S017H3–3 locus, and those originating from the resistant parent form three groups close to the S124E1–2 locus (Fig. 5). In addition, only one STS (Ha-NBS2/R) completely co-segregates with the *Pl* genes conferring resistance to races 100 and 300.

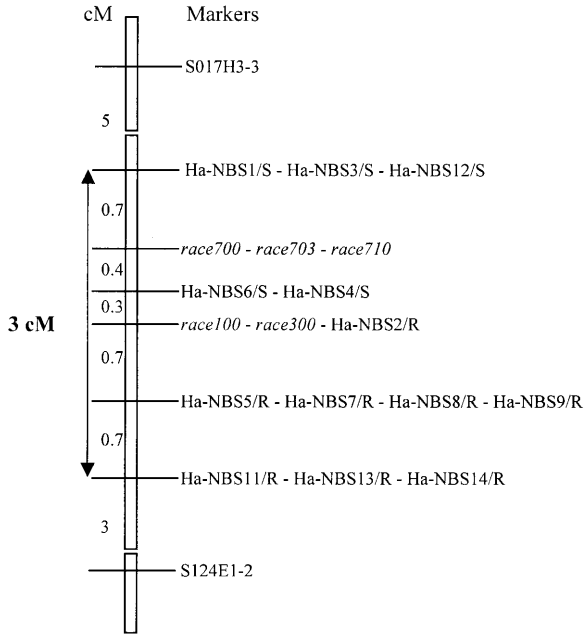

Fig. 5 Genetic map of the *Pl6* locus of the sunflower showing the localisation of 13 STS sequences. *Pl1* and *Pl2* correspond to genes conferring resistance to *P. halstedii* races 100 and 300 (formerly races 1 and 2). *race700*, *race703* and *race710*, correspond to uncharacterised *Pl* genes conferring resistance to *P. halstedii* races 700, 703 and 710 respectively (formerly races C, B and A). S017H3–3 and S124E1–2 are the closest RFLP markers described in Gentzbittel et al. (1999)

Discussion

The *Pl6* locus contains sequences belonging to the TIR-NBS-LRR class

Since the initial cloning of some plant resistance genes, several research groups have demonstrated that PCR-amplification of conserved disease-resistance motifs can be used to identify disease-resistance loci (Kanazin et al. 1996; Leister et al. 1996; Yu et al. 1996; Hayes and Saghai Maroof 2000). However, difficulty remains in identifying the functional resistance genes; there are often sequences that show great similarity to resistance genes but do not code for a functional product. Degenerate PCR was used to detect a major locus for resistance to downy mildew in sunflower (Gentzbittel et al. 1998). To clone *Pl* genes conferring resistance to *P. halstedii*, we needed to develop additional markers tightly linked to these genes. This is possible using either techniques such as AFLP or targeted strategies such as the modified AFLP method described by Hayes and Saghai Maroof (2000). In the present study, we targeted the resistance loci using PCR and specific primers in two ways: first, we cloned and sequenced a full-length RGA using RACE-PCR and specific primers; then we used Bulked Segregant Analysis (Michelmore et al. 1991) and specific primers derived from one complete RGA cloned from sunflower (accession number, AF316405). Thirteen

sequence-tagged-sites or STSs (accession numbers, see Table 2) within the *Pl*6 locus were then developed.

Sequencing of the full-length sunflower RGA identified complete signature domains of plant R-genes: TIR, NBS, LRR. The combination of these domains placed this sunflower RGA in the *L6* (Lawrence et al. 1995), *N* (Whitham et al. 1994), and *RPP1* (Botella et al. 1998) family of plant resistance genes. The cloned STS sequences showed 22–48% amino-acid identity with cloned R-genes (*N* and *L6*). The 13 STSs differed significantly with percent identities varying between 57 and 98%, except in the case of Ha-NBS11 which showed lower percent identities with the other STSs, varying from 25 to 46%. The variability observed was due both to substitutions and to insertions-deletions (Fig. 3). For example, Ha-NBS2/R shows a deletion consisting of 16 amino-acids. Ha-NBS7/R and Ha-NBS14/R show a 25 amino-acid deletion at the same position (Fig. 3) which may account for their grouping in the phylogenetic analysis and suggests that they may have arisen from the same mutational event. Sequence comparison of the 13 STSs showed that they contain some highly conserved domains, in agreement with the finding of Noël et al. (1999), who showed that the predicted proteins of the 12 non-truncated members of the *RPP5* family in *A. thaliana* share a high level of sequence conservation within the TIR and the NBS domains. These authors concluded that this is consistent with the TIR domain being involved as a part of the effector portion of the protein, and the NBS being a conserved pocket for the binding of ATP or GTP.

The *Pl6* locus may contain functional and non-functional TIR-NBS-LRR sequences

The preliminary studies on the expression of the STSs showed that either these STSs belong to nonfunctional genes, or that the level of transcription of these genes is very low and was not detected by traditional RT-PCR methods. Another possibility is that these STSs belong to R-genes which are induced by infection. For example, *Xa1* mRNA was detected from rice leaves at 5 days after inoculation of both compatible and incompatible strains of *Xanthomonas oryzae* pv *orysae*, but was not detected in intact leaves (Yoshimura et al. 1998). With the primer HaP3, the fragment obtained of approximately 700 bases may correspond to the Ha-NBS11 STS from which two introns (399 bp and 615 bp) had been removed. However, these results are preliminary and require other investigations, in particular the production of complete cDNA.

The *Pl6* locus may span several megabases

In this study, the 13 STS markers were all mapped to the *Pl6* resistance locus on linkage group 1 of the RFLP map described by Gentzbittel et al. (1999), indicating that either this subfamily of sequences is highly clustered and/or that the primers used are highly specific to this region. Recently, a second major locus for resistance to downy mildew in sunflower has been detected (Bert et al. 2001) and we are trying different primer combinations to test whether this second locus contains TIR-NBS-LRR sequences. In addition, the genetic mapping of the 13 STSs appears to support the phylogenetic analysis, with the "susceptible" STSs clustered in one region while the "resistant" STSs are clustered in another region of the locus (Fig 4).

These STS markers were located within a genetic distance of about 3 cM, which suggests that either the *Pl6* locus exhibits a high degree of recombination and/or that it is very large and complex. It is possible that this recombination distance reflects a physical organisation of a cluster probably extending over several hundreds of kilobases, even some megabases. Many plant resistance genes appear to be organised as complex clusters. Only a few clusters of resistance genes have been sequenced. The complete sequencing of the *RPP5* cluster in *Arabidopsis*, the *Cf-4/9* and *Pto* clusters in tomato, and partial sequencing of the *Dm3* cluster in lettuce revealed highly duplicated regions containing more than 24 resistance-gene homologs (Michelmore 2000). The *Dm3* locus of lettuce is one example in which numerous related copies of resistance gene homologues are spread over several megabases of one chromosome (Meyers et al. 1998a, b). It will be interesting to develop a physical map of the *Pl6* locus and to see whether it is as complex as the *Dm3* locus in lettuce.

The STSs should facilitate marker-assisted selection

Numerous molecular markers closely linked to resistance genes have been recently developed in many crops; for example, near the *Xa21* gene in rice (Williams et al. 1996), the *N* gene homologs in potato (Hehl et al. 1999), the *Rsv1* gene in soybean (Hayes and Saghai Maroof, 2000), or the *Rph*7.g locus in barley (Brunner et al. 2000). In sunflower, the availability of 13 specific PCR-based markers for the *Pl6* locus should facilitate the selection and introgression of these resistance genes into new varieties. In our laboratory, the primer pair HaP1 has been used successfully for the introgression of the *Pl6* locus into a completely susceptible line (unpublished results). When the STSs were used as probes in Southern analysis, they all gave complex patterns. In addition, the first three primer pairs used in this study gave polymorphism not only between the two parents but also between a set of 24 different sunflower lines (unpublished results). Together, these results suggest that there is a potential to develop other molecular markers for the *Pl6* locus using other primer pairs specific, for example, to the LRR domain.

Acknowledgements This study was supported by PROMOSOL, CETIOM and a grant "Contrat de Plan Etat-Région Semences et Plants".

References

Altschul SF, Madden TL, Shaffer 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

Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL (2000) The Pfam protein families database. Nucleic Acids Res 28:263–266

Bert PF, Tourvieille de Labrouhe D, Philippon J, Mouzeyar S, Jouan I, Nicolas P, Vear F (2001) Identification of a second group of genes controlling resistance to downy mildew (*Plasmopara halstedii*) in sunflower (*Helianthus annuus L*). Theor Appl Genet (in press)

Bogorad L, Gubbins EJ, Krebbers E, Larrinua IM, Mulligan BJ, Muskavitch KMT, Orr EA, Rodermel SR, Schantz R, Steinmetz AA, de Vos G, Ye YK (1983) Cloning and physical mapping of maize plastid genes. Methods Enzymol 97:524–554

Botella MA, Parker JE, Frost LN, Bittner-Eddy PD, Beynon JL, Daniels MJ, Holub EB, Jones JDG (1998) Three genes of the *Arabidopsis RPP1* complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. Plant Cell 10:1847–1860

Brunner S, Keller B, Feuillet C (2000) Molecular mapping of the *Rph*7.g leaf rust resistance gene in barley (*Hordeum vulgare* L.). Theor Appl Genet 101:783–788

Burge C, Karlin S (1997) Prediction of complete gene structures in human genomic DNA. J Mol Biol 268:78–94

Ellis JG, Jones D (1998) Structure and function of proteins controlling strain-specific pathogen resistance in plants. Curr Opin Plant Biol 1:288–293

Gentzbittel L, Mouzeyar S, Badaoui S, Mestrie E, Vear F, Tourvieille de Labrouhe D, Nicolas P (1998) Cloning and molecular markers for disease resistance in sunflower, *Helianthus annuus* L. Theor Appl Genet 96:519–525

Gentzbittel L, Mestries E, Mouzeyar S, Mazeyrat F, Badaoui S, Vear F, Tourvieille de Labrouhe D, Nicolas P (1999) A composite map of expressed sequences and phenotypic traits of the sunflower (*Helianthus annuus* L.) genome. Theor Appl Genet 99:218–234

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

Hehl R, Faurie E, Hesselbach J, Salamini F, Whitham S, Baker B, Gebhardt C (1999) TMV resistance gene *N* homologues are linked to *Synchytrium endobioticum* resistance in potato. Theor Appl Genet 98:379–386

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 MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

Lawrence GJ, Finnegan EJ, Ayliffe MA, Ellis JG (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

Leister D, Ballvora A, Salamini F, Gephardt C (1996) A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. Nature Genet 14:421–428

Martin GB (1999) Functional analysis of plant disease resistance genes and their downstream effectors. Curr Opin Plant Biol 2:273–279

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 and 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 (2000) Genomic approaches to plant disease resistance. Curr Opin Plant Biol 3:125–131

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomes. Proc Natl Acad Sci USA 88:9828–32

Miller JF, Gulya TJ (1991) Inheritance of resistance to race 4 of downy mildew derived from interspecific crosses in sunflower. Crop Sci 31:40–43

Mindrinos M, Katagiri F, Yu GL, Ausubel FM (1994) The *A. thaliana* disease resistance gene *RPS*2 encodes a protein containing a nucleotide binding site and leucine-rich repeats. Cell 78: 1089–1099

Nicholas KB, Nicholas HB, Deerfield, DW (1997) EMBnet News 4:14

Noël L, Moores TL, van der Biezen EA, Parniske, Daniels MJ, Parker JE, Jones JDG (1999) Pronounced intraspecific haplotype divergence at the *RPP5* complex disease resistance locus of *Arabidopsis*. Plant Cell 11:2099–2112

Saghai Maroof MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc Natl Acad Sci USA 81:8014–8018

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgens DG (1997) The CLUSTAL-X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Vear F, Gentzbittel L, Philippon J, Mouzeyar S, Mestries E, Roeckel-Drevet P, Tourvieille De Labrouhe D, Nicolas P (1997) The genetics of resistance to five races of downy mildew (*Plasmopara halstedii*) in sunflower (*Helianthus annuus*). Theor Appl Genet 95:584–589

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

Williams CE, Wang B, Holsten TE, Scambray J, de Assis Goes da Silva F, Ronald PC (1996) Markers for selection of the rice *Xa*21 disease resistance gene. Theor Appl Genet 93:1119–1122

Yoshimura S, Yamanouchi U, Katayose Y, Toki S, Wang ZX, Kono I, Kurata N, Yano M, Iwata N, Sasaki T (1998) Expression of Xa1, a bacterial blight-resistance gene in rice, is induced by bacterial inoculation. Proc Natl Acad Sci USA 95:1663–1668

Yu YG, Buss GR, 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