Theor Appl Genet (2004) 109:176–185
DOI 10.1007/s00122-004-1613-0

ORIGINAL PAPER

O. Radwan · M. F. Bouzidi · P. Nicolas · S. Mouzeyar

Development of PCR markers for the *Pl5/Pl8* locus for resistance to *Plasmopara halstedii* in sunflower, *Helianthus annuus* L. from complete CC-NBS-LRR sequences

Received: 9 December 2003 / Accepted: 28 January 2004 / Published online: 9 March 2004

Abstract Sunflower downy mildew, caused by *Plasmopara halstedii*, is one of the major diseases of this crop. Development of elite sunflower lines resistant to different races of this oomycete seems to be the most efficient method to limit downy mildew damage. At least two different gene clusters conferring resistance to different races of *P. halstedii* have been described. In this work we report the cloning and mapping of two full-length resistance gene analogs (RGA) belonging to the CC-NBC-LRR class of plant resistance genes. The two se-quences were then used to develop 14 sequence tagged sites (STS) within the *Pl5/Pl8* locus conferring resistance to a wide range of *P. halstedii* races. These STSs will be useful in marker-assisted selection programs.

Introduction

In "gene-for-gene" interactions between plants and their pathogens, the reaction of incompatibility, i.e. no disease, requires a resistance (*R*) gene in the plant, and a corresponding avirulence (*Avr*) gene in the pathogen. In this system, the *R* genes are presumed to enable plants to detect *Avr*-gene-specified pathogen molecules, and to initiate signal transduction to activate defenses (Hammond-Kosack and Parker 2003).

These *R* gene products can be grouped into five classes based on their structural features; the vast majority are characterized by the presence of leucine-rich repeat (LRR) motifs. LRR-containing R proteins can be further subdivided into two classes based on additional structural features. The first class includes proteins with a predicted hydrophobic membrane-anchoring domain and a predicted extra-cytosolic N-terminal LRR motif (Jones et al. 1994; Song et al. 1995; Bent 1996). The second class of LRR-containing R proteins can be distinguished by the presence of a putative tripartite nucleotide binding site (NBS) N-terminal to the LRR domain (Staskawicz et al. 1995; Bent 1996). The NBS-LRR class of R proteins represents by far the majority of functionally described *R* genes. These genes contain at least three discernible regions: a variable N terminus, a nucleotide binding site, and leucine-rich repeats. Two kinds of N termini are present in NBS-LRR proteins. The first contains coiled coils (CC) that are thought to play a role in protein-protein interactions. A CC motif appears in the N terminus of NBS-LRR proteins from both monocotyledons and dicotyledons (Meyers et al. 1999; Pan et al. 2000) including *RPM1* (Grant et al. 1995), *RPS2* (Bent et al. 1994; Mindrinos et al. 1994) from *Arabidopsis*, *M1* from tomato (Milligan et al. 1998) *R1* from potato (Ballvora et al. 2002) and *rp3* from maize (Webb et al. 2002). The other type of N terminus shows homology with *Drosophila* Toll or human interleukin receptor-like protein (TIR). This second type is absent from the NBS-LRR proteins of monocots (Meyers et al. 1999) and appears in dicot resistance proteins including *L6* from flax (Lawrence et al. 1995), *N* from tobacco (Whitham et al. 1994) and *RPP5* from *Arabidopsis thaliana* (Parker et al. 1996).

Downy mildew, caused by *Plasmopara halstedii* (Farl.) Berl. & de Toni, is one of the main diseases causing economic losses in cultivated sunflower (*Helianthus annuus* L.). The major dominant genes denoted by *Pl* confer resistance to this disease, following a pattern which agrees well with the gene-for-gene hypothesis of Flor (1955). Genetic studies showed that the *Pl6* locus from wild *H. annuus* (Miller and Gulya 1991) 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 clustering of *P. halstedii* resistance genes in sunflower. A second region carrying downy mildew resistance genes was reported by

Communicated by C. Möllers

O. Radwan · M. F. Bouzidi · 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
Tel.: +33-4-73407911
Fax: +33-4-73407914

Bert et al. (2001), who mapped the *Pl5* locus from the Russian population Progress to linkage group 6 of the same sunflower map.

Radwan et al. (2003) cloned and sequenced 16 RGA based on the sequences obtained by Gedil et al. (2001). The genetic mapping of these RGAs showed that the non-TIR-NBS-LRR RGAs were all clustered, and linked to the *Pl5*/*Pl8* locus on linkage group 6 of the map of Gentzbittel et al. (1999). However the sequences obtained were only partial (e.g. 300 bp) and further characterization of the *Pl5/Pl8* locus was needed.

In this paper, we report the cloning, sequencing and mapping of full-length sequences belonging to the CC-NBS-LRR class of resistance genes in sunflower. These sequences were also used to develop specific PCR-based markers for the *Pl5/Pl8* locus. The use of these new PCR based markers in a marker-assisted selection program is proposed.

Materials and methods

Sunflower genotypes and resistance tests

The genotypes and the phenotypic segregation for resistance to downy mildew were described previously (Radwan et al. 2003). The resistance tests for downy mildew were carried out as described by Mouzeyar et al. (1994).

DNA and RNA manipulations

Young leaf tissue from the F_2 plants was collected and freeze-dried. DNA was isolated using the CTAB method, as described by Saghai Maroof et al. (1984). Equal quantities of DNA were bulked from 12 homozygous-resistant and from 12 homozygous-susceptible F_2 plants, according to the bulked segregant analysis method (Michelmore et al. 1991) to give the two DNA bulks of each cross. RNA was extracted from healthy 15-day-old cotyledons of OC, YSQ, CAY, and QIR8, by using the method described by Bogorad et al. (1983). Poly $(A)^+$ mRNA was isolated by the PolyAtract mRNA Isolation system (Promega).

Rapid amplification of cDNA ends of two partial RGAs

Five prime and 3′ ends of cDNA were obtained by rapid amplification of cDNA ends (RACE) using the "Marathon cDNA amplification kit" (Clontech, Ozyme France). Gene-specific primers and nested gene-specific primers were designed based on the sequences of the 248- and 277-bp RGA products (NTIR11, accession number AF528547; NTIR3, accession number AF528539) obtained by Radwan et al. (2003). The Ha-PA and Ha-PC primer pairs were used to amplify the 5′ and 3′ ends of the Ha-NTIR11 cDNA, respectively, and then the Ha-PB and Ha-PD primer pairs were used for the nested PCR. To amplify the 5′ and 3′ ends of the Ha-NTIR3 cDNA, the Ha-PE and Ha-PG primer pairs were used for the first PCR step, then the Ha-PF and Ha-PH primer pairs for the nested PCR. The names and sequences of these primers are given Table 1. One microlitre of cDNA was used as a template for the first PCR round, then the PCR products were diluted 3/100 and 10 μl were used as a template for the nested PCR. PCR reactions (50 μl) contained 1 U (1 μl) of *Taq* DNA polymerase (Advantage 2, Clontech, France), 1×*Taq* polymerase buffer [40 mM Tricine-KOH pH 8.7, 15 mM KOAc and 3.5 mM $Mg(OAc)_2$], 0.5 mM of each dNTP and 1 μM of each primer. PCR was carried out in a 2400 Perkin-Elmer thermocycler under the following conditions: for the first PCR, an initial denaturation at 94°C for 3 min was followed by 40 cycles of 94°C for 30 s, 65°C for 60 s and 72°C for 2.5 min, and for the nested PCR, an initial denaturation at 94°C for 3 min, 40 cycles of 94°C for 5 s, 68°C for 4 min. The amplified products were cloned into the pGEM-T Easy vector (Promega) and sequenced by Genome Express (Grenoble, France).

Amplification of full-length cDNAs and genomic sequences of Ha-NTIR11 and Ha-NTIR3 RGAs

To amplify the corresponding full-length cDNAs of the two partial RGAs, we selected specific primers based on the sequences of the 5′ and 3′ RACE-PCRs (accession nos. AY490792, AY490795, AY490796 and AY490794). The forward primers contained the initiation codon ATG and the reverse primers were designed prior to the poly $(A)^+$ tail. To amplify the complete sequence of Ha-NTIR11, we used the Ha-PI primer pair for the first PCR and then the Ha-PJ primer pair for the nested PCR. The amplification of the complete sequence of Ha-NTIR3 was carried out using the Ha-PK primer pair for the first PCR and then the Ha-PL primer pair for the nested PCR. These primers are given Table 1. One hundred nanograms of genomic DNA or 1 μl of cDNA were used as templates for the first round PCR. The PCR products were then diluted 3/100 and 10 μl were used as templates for the nested PCR. PCR was carried under the following conditions: initial denaturation at 95°C for 3 min, 40 cycles of 95°C for 10 s, 60°C for 30 s and 72°C for 6 min for the first PCR, and initial denaturation at 95°C for 3 min,

Table 1 Forward and reverse sequence primers that were used to amplify the 5′ and 3′ ends and the full length Ha-NTIR11g and Ha-NTIR3A RGAs

Primer pair	Forward primer sequences	Reverse primer sequences
Ha-PA	5′CCATCCTAATACGACTCACTATAGGGC3′[a]	5′CCCAGTCGTCATATGTTTCATTCC3′
Ha-PB	5′ACTCACTATAGGGCTCGAGCGGC3′[b]	5′GCCCTCAAGTTTCTCTTTAAGAGC3′
Ha-PC	5′CGCGAGTGAAGGGTCAGTTTGAAC3′	5′CCATCCTAATACGACTCACTATAGGGC3′[a]
Ha-PD	5′TCATGGCATGGGTTTGCGTGTCCG3′	5′ACTCACTATAGGGCTCGAGCGGC3′[b]
Ha-PE	5′CCATCCTAATACGACTCACTATAGGGC3′[a]	5′CAGAAACACAAACCCATGTCTTGGG3′
Ha-PF	5′ACTCACTATAGGGCTCGAGCGGC3′[b]	5′GTTCAAAGTGATCCTTCACTTGTGCG3′
Ha-PG	5′CACAAGTGAAGGATCACTTTGAAC3′	5′CCATCCTAATACGACTCACTATAGGGC3′[a]
Ha-PH	5′CCAAGACATGGGTTTGTGTTTCTG3′	5′ACTCACTATAGGGCTCGAGCGGC3′[b]
Ha-PI	5′CCTCTTCACTGTTAGTTAACCATGG3′	5′TTACACTTAACGGCTTGACCCAAG3′
Ha-PJ	5′ATGGCTGAAACCGCTGTTACTGCCC3′	5′GACCCAAGAAGCTATGGGGTCAAG3′
Ha-PK	5′GTTAACCATGGCTGATGAAACTCTTGC3′	5′CCTCTGGTCTATTTTGATTTTGGGG3′
Ha-PL	5′CCATGGCTGAAACTCTTGCAAATG3′	5′CAGCGTCTCTGGTAGATCGTTCACC3′

[a] Adaptor primer 1 (AP1)
[b] Adaptor primer 2 (AP2)

Table 2 Forward and reverse sequence primers that were used to amplify the STSs. Their locations within the genomic DNA sequence of Ha-NTIR11g (accession number AY490793) and the cDNA sequence of Ha-NTIR3A (accession number AY490791) are also indicated

Primer pair	Forward primer sequences	Reverse primer sequences
Ha-P1[a]	5′GCCCAAAATTGAAAGAAAGGTGTG3′ Nucleotides 3,752 to 3,775	5′GGCGAAATTGGTTCCCGTGAGTCG3′ Nucleotides 6,111 to 6,088
Ha-P2	5′AATCTTGAGTCATTACCCGAGC3′ Nucleotides 3,360 to 3,381	5′CAGCGTCTCTGGTAGATCGTTCACC3′ Nucleotides 3,866 to 3,842
Ha-P3[a]	5′TAGTTAACCATGGCTGAAACCGCTG3′ Nucleotides –9 to 16	5′TTTGAAAGATAAGTTCGCCTCTCG3′ Nucleotides 2,169 to 2,146
Ha-P4[a]	5′GCTGTTACTGCCCTCTTCAAAGTC3′ Nucleotides 13 to 36	5′CCCAACTCGACATATCTTCAAACC3′ Nucleotides 2,446 to 2,423
Ha-P5[a]	5′TAGTTAACCATGGCTGAAACCGCTG3′ Nucleotides –9 to 16	5′CCCCATATTGACAAAGAGTTGAGG3′ Nucleotides 3,116 to 3,093
Ha-P6[a]	5′TAGTTAACCATGGCTGAAACCGCTG3′ Nucleotides –9 to 12	5′CGTCTCTGGTAGATCGTTCACCTT3 Nucleotides 3,714 to 3,691

[a] Primers that were selected from the Ha-NTIR11g sequence

35 cycles of 95°C for 5 s, 68°C for 30 s and 72°C for 6 min for the nested PCR. The amplified products were cloned and sequenced as described previously.

Southern blotting

DNA digestion and Southern hybridization were performed as described previously (Gentzbittel et al. 1999) using two restriction enzymes: *Eco*R1 and *Eco*R5. The polymorphic loci were scored on 150 F_2 plants from the CAY×QIR8 and OC×YSQ crosses. The probes were prepared as described previously (Radwan et al. 2003).

Amplification and cloning of polymorphic PCR fragments

Six specific primer pairs (Table 2) were chosen randomly, based on the complete sequence of Ha-NTIR11g (accession no. AY490793) and Ha-NTIR3A (accession no. AY490791) RGAs, and tested for polymorphism using the bulked segregant analysis method (Michelmore et al. 1991). The PCR reactions (25 μl) contained 50 ng of sunflower DNA, 0.2 mM of each dNTP, 0.4 U (0.4 μl) of *Taq* DNA polymerase (Advantage 2, Clontech, France), 1×*Taq* polymerase buffer and 0.4 μM of each primer. PCR was carried under the following conditions: initial denaturation at 95°C for 3 min, 35 cycles of 95°C for 30 s, 60°C for 30 s and 72°C for 2 min. 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 as described previously.

DNA sequencing and sequence analysis

Two clones of each 5′ and 3′ RACE-PCR, one clone of each full-length sequence and one clone for each STS were chosen randomly and completely sequenced on both strands using the Dye-Terminator method (Genome Express, France). The nucleotide and amino acid sequences were compared with those released in the GenBank databases using the BLAST analysis program (Altschul et al. 1997). The sequences were aligned using the CLUSTAL X software with default options (Thompson et al. 1997) and the resulting alignments were shaded using the GENEDOC software (Nicholas et al. 1997).

Full-length RGA and STS mapping

Marker order and genetic distances were calculated using MAPMAKER 3.b software (Lander et al. 1987). Markers were ordered with a LOD value threshold of 3.0 and a maximum recombination fraction of 50. The polymorphic loci detected with Ha-NTIR11g and Ha-NTIR3A RGAs and the STSs amplified with the primer pair Ha-P1 were scored as co-dominant markers. The remaining STSs were scored as dominant markers. The polymorphic loci were mapped using 150 F_2 individuals from the two crosses OC×YSQ and CAY×QIR8. Other polymorphic markers described by Radwan et al. (2003) were included in this study and two genetic maps were constructed for linkage group 6 of the map of Gentzbittel et al. (1999).

Results

Isolation and cloning of the 5′ and 3′ ends of the two RGAs

The partial sequences of the two RGAs, Ha-NTIR11 and Ha-NTIR3 (Radwan et al. 2003), were used as templates to isolate the remaining 5′ and 3′ portions of the corresponding cDNAs. Two clones for each of the 5′ and 3′ ends of Ha-NTIR11 were randomly selected and sequenced. The two clones of the 5′ end were the same length, 858 bp, (accession no. AY490792). The sequences of the two 3′ end clones (accession no. AY490796) included a poly $(A)^+$ tail. For sequencing the 5′ and 3′ ends of Ha-NTIR3, two clones of each end were selected. Again, the two 5′ clones were the same length, 673 bp (accession no. AY490795), while the partial sequences of the two 3′ end clones (accession no. AY490794) indicated that there are five putative stop codons at positions 416, 324, 305, 9, and 3 before the poly $(A)^+$ tail.

Amplification of the full-length cDNA and genomic sequences

The sequences of 5′ and 3′ ends of the RACE-PCR products were used to design specific primers in order to amplify the full-length cDNA and genomic sequences. We used nested PCR to obtain the full-length sequence of Ha-NTIR11. A single band was produced, which was cloned and sequenced; the genomic sequence length was 6,780 bp (accession number AY490793) corresponding to a cDNA of 5,154 bp. This sequence contains one ORF corresponding to an exon of 3848 bp and was denoted Ha-

NTIR11g. We used the same approach to obtain the full-length sequence of Ha-NTIR3. Here, the nested PCR produced a single band from either the cDNA or gDNA templates. The band was cloned and two clones, denoted Ha-NTIR3A and Ha-NTIR3B, were randomly selected and sequenced. The lengths of the two clones are 4,034 bp (accession no. AY490791) and 3,986 bp (accession no. AY490797) with a putative stop codon at position 3,986 of Ha-NTIR3A and 3,861 of Ha-NTIR3B.

Sequence analysis

The predicted protein structures of the RGA clones Ha-NTIR11g and Ha-NTIR3A determined using the PFAM (http://pfam.wustl.edu) and SMART (http://smart.embl-heidelberg.de) databases were 1,279 and 1,302 amino acids long respectively, and showed similarity to other resistance genes. They shared 53% and 58% identity with resistance protein candidate (*RGC 1b*) from *Lactuca sativa* (Shen et al. 1998), 32% and 35% identity with the *RPP13*-like protein encoded by the downy mildew resistance gene from *Arabidopsis thaliana* (Sato et al. 2000), and 30% and 32% identity with the I2C-2 protein encoded by the tomato wilt resistance gene from *Lycopersicon esculentum* (Ori et al. 1997). Both have 15% identity with the RPM1 protein encoded by the *Arabidopsis* bacterial wilt resistance gene (Grant et al. 1995), 13% identity with the L6 protein encoded by the flax rust resistance gene from *Linum usitatissimum* (Lawrence et al. 1995), and 9% and 8% identity with the PU3 protein encoded by the downy mildew resistance gene candidate from *Helianthus annuus* (Bouzidi et al. 2002). The amino acids sequences predicted from Ha-NTIR11g and Ha-NTIR3A were divided into three domains (Fig. 1). The first is a CC domain (amino acids 1–154 and 1–156 for the Ha-NTIR11g and Ha-NTIR3A proteins respectively). This domain contains a leucine zipper (LZ) (amino acids 26–57) for the two proteins. The second domain is an NBS (amino acids 155–542 and 157–546 for Ha-NTIR11g and Ha-NTIR3A, respectively). This domain includes the P-loop (Kinase 1-a), Kinase 2 and Kinase 3-a (RNBS-B) motifs of the NBS (Aravind et al. 1999) and the RNBS-A, RNBS-C, GLPL, RNBS-D and MHD motifs that are conserved in NBS-LRR proteins (van der Biezen and Jones 1998; Meyers et al. 1999). The third domain is a LRR and the C terminus (amino acids 542–1279 for Ha-NTIR11g and 547–1302 for Ha-NTIR3A). Several LRR motifs were detected in this region, however, most of them are imperfect except for three motifs in Ha-NTIR11g (amino acids 604–622, 623–644 and 1054–1078) and four motifs in Ha-NTIR3A (amino acids 604–626, 627–650, 1109–1131 and 1149–1171). Furthermore, amino acid sequences (LLRVLSL) and (LLGVLSL) in a second LRR for Ha-NTIR3A and Ha-NTIR11g respectively are conserved in other LZ NBS-LRR type R proteins (LLRVLDL), suggesting functional significance (Bittner-Eddy et al. 2000). The overall identity between the two proteins is 58%, whereas the percent identity of the CC domain 53%, of the NBS domain 68% and of the LRR/C terminus 54%.

Southern hybridization

Southern hybridization analysis was carried out to detect the genomic organization of Ha-NTIR11g and Ha-NTIR3A. Southern blots showed multiple bands of varying intensity in each enzyme restriction digestion and these bands represented Ha-NTIR11g and Ha-NTIR3A or their homologous sequences in sunflower. Polymorphic bands were detected between the susceptible and resistant parents CAY and QIR8 for Ha-NTIR11g and OC and YSQ for Ha-NTIR3A in all enzyme restriction digestion used. Two polymorphic loci with *Eco*RI and *Eco*R5-digestions for CAY/QIR8 and OC/YSQ respectively were selected and then scored on 150 F_2 plants of each cross. The two loci were mapped to linkage group 6 (Fig. 2).

Amplification and cloning of sequence tagged sites

Five primer pairs (Table 2) were selected from the complete sequence of Ha-NTIR11g and one primer pair from that of Ha-NTIR3A, then tested for their abilities to amplify polymorphic fragments capable of distinguishing between the susceptible sunflower lines CAY and OC and the resistant lines QIR8 and YSQ. Bulked segregant analysis (Michelmore et al. 1991) was used to detect PCR markers potentially linked to the *Pl5*/*Pl8* locus (Fig. 3a–h). Amplification from DNA of the parental lines CAY and QIR8 with the primer pair Ha-NTP3 gave one major polymorphic band that differed between the two parents and the bulks (Fig. 3a). The same band was detected in OC and YSQ (Fig. 3b). The primer pair Ha-NTP4 amplified one band that differed between the two parents CAY and QIR8 and the bulks (Fig. 3c). However the same primer pair failed to amplify any polymorphic bands from the two parents OC and YSQ. The Ha-NTP5 and Ha-NTP6 primer pairs each gave a single band in the two parents CAY and QIR8 and the bulks (Fig. 3d, e) but they failed to amplify any polymorphic bands that differed between the two parents OC and YSQ. The Ha-NTP1 primer pair gave five polymorphic bands in the two parents CAY and QIR8 and the bulks (Fig. 3f). The Ha-NTP1 primer pair gave three polymorphic bands in the two parents OC and YSQ and the bulks (Fig. 3g). The primer pair Ha-NTP2 gave one polymorphic band differing between the two parents OC and YSQ and the bulks (Fig. 3h). This primer pair was not tested on the other parents CAY and QIR8. All of the polymorphic bands were individually cloned and sequenced. The sequence length and the origin of each STS are given in Table 3.

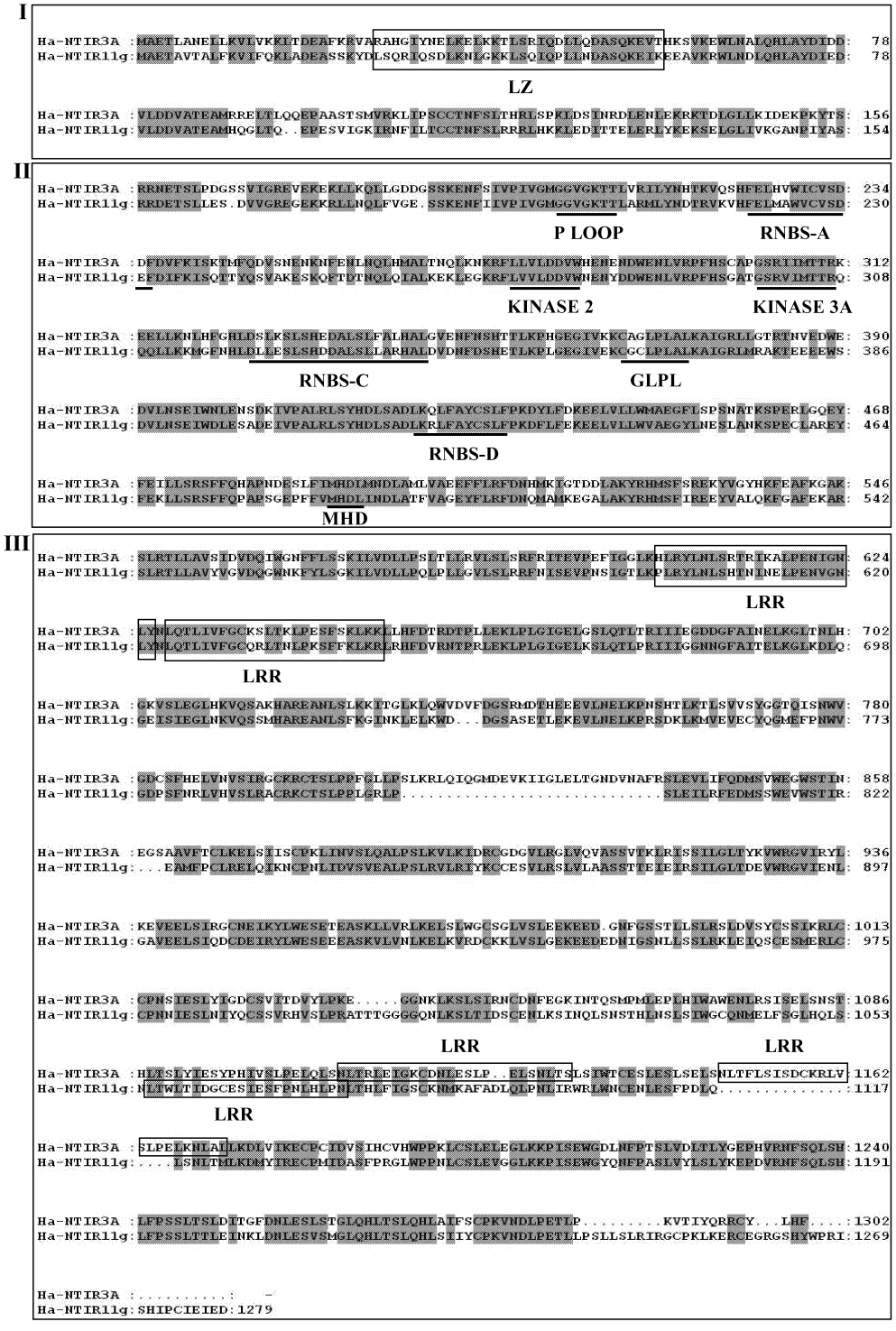

Fig. 1 Complete alignment of the deduced amino acid sequences of two RGAs (Ha-NTIR11g and Ha-NTIR3A, accession numbers AY490793 and AY490791 respectively) that belong to the *Pl5*/*Pl8* locus on sunflower linkage group 6. CLUSTAL X was used for the alignment. Identical residues between the two sequences (58%) are *shaded* using GENDOC software. The sequence is divided into three domains: *I* indicates the CC domain, *II* indicates the NBS domain and *III* the LRR domain. Each domain contains the motifs that are indicated. The different motifs were identified according to Meyers et al. (1999)

Sequence analysis and comparison of the 14 STSs

Due to the position of the primers used, we expected that the amplifications using different primer pairs would give different product sizes. The 14 STSs that we identified were classified into four groups according to their homology with other resistance genes in the genetic databases. The first class includes Ha-NT8R3, Ha-NT5R1 and Ha-NT8R4, the second includes Ha-NT5S3, the third includes Ha-NT8R5 and Ha-NT8R6, and the fourth includes Ha-NT8S1, Ha-NT8S2, Ha-NT8R1, Ha-NT8R2, Ha-NT8R7, Ha-NT5S1, Ha-NT5S2 and Ha-NT5R2.

The first class showed homology with the first region of other resistance genes and shared identity with *RGC 1b* from *Lactuca sativa* (Shen et al. 1998), *RPP13* from *Arabidopsis thaliana* (Sato et al. 2000) and *I2C-2* from

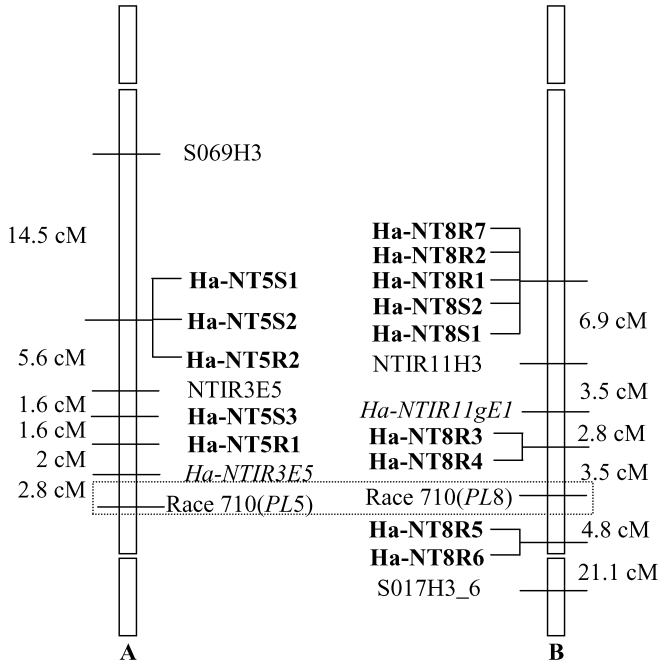

Fig. 2A, B Genetic map of the *Pl5* (**A**) / *Pl8* (**B**) locus on sunflower linkage group 6 showing localization of two polymorphic loci after *Eco*R1 and *Eco*R5 digestion (shown in *italics*) that are detected when Ha-NTIR11g and Ha-NTIR3A are used as probes. Fourteen STSs, shown in *bold*, were mapped also. Two RFLP markers (S069H3 and S017H3_6) and two partial RGA sequences (NTIR11H3 and NTIR3E5; Radwan et al. 2003) are shown. Genetic distances were calculated using the Kosambi mapping function and are shown in centimorgans (cM). The suffixes E1, E5 and H3 indicate the restriction enzymes *Eco*RI, *Eco*RV and *Hind*III respectively

Lycopersicon esculentum (Ori et al. 1997) with percentage identities of 56–58%, 36% and 36% respectively. The second class showed homology with the LRR region of other resistance genes. The most closely related plant resistance genes were *rp3* from *Zea mays* (Webb et al. 2002), *I2C-3* from *L. esculentum* (Ori et al. 1997) and *I2C-5* from *Lycopersicon pimpinellifolium* (Sela-Burrlage et al. 2001) with percentage identities of 36%, 36% and 34%, respectively. The third class of STS did not share any homology with the other resistance genes although they showed homology with other gene products such as poly-protein from *Oryza sativa* (accession number BAB90375), with a percentage identity of 40%. The fourth class of STS interrupts the 3′ UTR of the Ha-NTIR11g RGA sequence and did not show homology with other resistance genes except for about 78–120 nucleotides (prior to the stop codon) that shared weak homology with other resistance gene proteins.

Mapping of the two complete RGAs and the 14 STS markers

To map the polymorphic loci corresponding to the complete RGA Ha-NTIR11g and Ha-NTIR3A, and the 14 STS markers, two genetic maps were constructed using two F_2 segregating populations. The first map (Fig. 2A) includes the *Pl5* locus, five STS markers, the Ha-NTIR3AE5 locus, the NTIR3E5 RGA and the S069H3 RFLP marker, while the second (Fig. 2B) contains the *Pl8* locus, nine STS markers, the Ha-NTIR11gE1 locus, the NTIR11H3 RGA and the S017H3_6 RFLP marker. The STSs that were amplified using the Ha-NTP3, Ha-NTP4, Ha-NTP5 and Ha-NTP6 primer pairs were used as dominant markers while the other STSs that were amplified using the Ha-NTP1 primer pair and the other markers were used as co-dominant markers. All the polymorphic loci we detected map to the distal region of linkage group 6 of the RFLP composite map developed by Gentzbittel et al. (1999). This linkage group corresponds to linkage group 13 of Yu et al. (2003). The Ha-NTIR3AE5 locus, Ha-NT8R3, Ha-NT5S3 and Ha-NT5S1/Ha-NT5S2/Ha-NT5R2 STSs mapped 2.8 cM, 4.8 cM, 6.4 cM and 13.6 cM from the *Pl5* locus. The Ha-NT8R5/Ha-NT8R6 STSs mapped 4.8 cM from the *Pl8* locus on one side while Ha-NT8R3/Ha-NT8R4, Ha-NTIR11gEI and Ha-NT8R1/Ha-NT8R2/Ha-NT8R7/Ha-NT8S1/Ha-NT8S2 mapped 3.5 cM, 6.3 cM and 16.7 cM from the other side of the *Pl8* locus. These STS markers are clustered within a genetic distance of about 13.6 cM and 21 cM for the *Pl5* and *Pl8* loci, respectively. Thus the *Pl5*/*Pl8* locus contains several copies of CC-NBS-LRR resistance gene analogs.

Discussion

Structure of encoded proteins and similarity to other resistance genes

The *Pl5* locus confers resistance to a wide range of downy mildew races but is susceptible to a US isolate of race 330, while the *Pl8* locus confers resistance to all known races of *Plasmopara halstedii*. These two loci have been mapped in the same area of linkage group 6 (Radwan et al. 2003). The predicted proteins of the full-length clones Ha-NTIR11g and Ha-NTIR3A are 1,279 and 1,302 amino acids long, respectively. Overall, the two predicted sequences share 58% identity and 71% similarity. Comparison with other available sequences of plant resistance genes revealed that these proteins belong to the nucleotide binding-LRR family of plant resistance genes. The closest similarity is to the putative protein encoded by lettuce *RGC 1b* (Shen et al. 1998). The two predicted RGA products carry a possible leucine zipper (LZ) (consensus XXXYXXL, where Y represents a hydrophobic residue) in their amino termini. It is proposed that this domain facilitates the formation of a coiled-coil (CC) structure to promote either dimerization or specific interactions with other proteins (Hammond-Kosack and Jones 1997). The CC structure appears in the N terminus of both monocotyledons and dicotyledons (Meyers et al. 1999; Pan et al. 2000) and includes *RPM1* (Grant et al. 1995), *RPS2* (Bent et al. 1994), *RPP13* (Bittner-Eddy 2000) from *Arabidopsis*; M1 from tomato (Milligan et al. 1998); R1

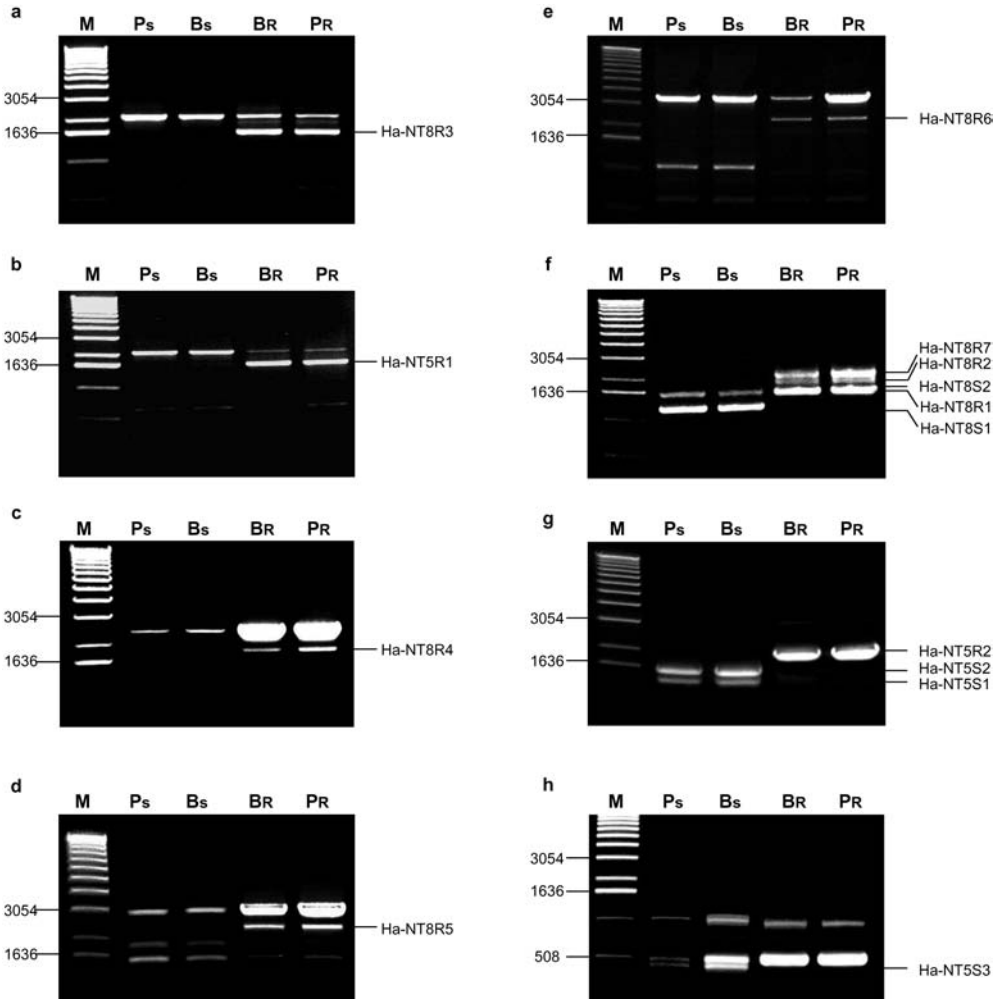

Fig. 3a–h STS amplification patterns. **a, g, h** Amplification patterns from the susceptible parent OC and the resistant parent YSQ. **b–f** Amplification patterns from the susceptible parent CAY and the resistant parent QIR8. *Lane M* 1 kb DNA ladder, P_S susceptible parent, B_S susceptible bulk, B_R resistant bulk, and P_R resistant parent. **a, b** STS amplification patterns with the Ha-P3 primer pair. **f, g** STS amplification patterns with the Ha-P1 primer pair. **c–e, h** Amplification patterns with the Ha-P4, Ha-P5, Ha-P6 and Ha-P2 primer pairs, respectively

Table 3 The PCR product sizes of the 14 STSs and the source of each. *S* Susceptible parent, *R* resistant parent. The corresponding accession numbers are given

STSs	Primer pairs	Parent	Locus	Length (bp)	Accession number
Ha-NT8R1	Ha-P1	R (QIR8)	*PL8*	1,569	BV097078
Ha-NT8R2	Ha-P1	R (QIR8)	*PL8*	2,119	BV097079
Ha-NT8R3	Ha-P3	R (QIR8)	*PL8*	1,584	BV097080
Ha-NT8R4	Ha-P4	R (QIR8)	*PL8*	1,840	BV097081
Ha-NT8R5	Ha-P5	R (QIR8)	*PL8*	2,419	BV097082
Ha-NT8R6	Ha-P6	R (QIR8)	*PL8*	2,437	BV097083
Ha-NT8R7	Ha-P1	R (QIR8)	*PL8*	2,237	BV097084
Ha-NT8S1	Ha-P1	S (CAY)	*PL8*	1,153	BV097074
Ha-NT8S2	Ha-P1	S (CAY)	*PL8*	1,610	BV097075
Ha-NT5R1	Ha-P3	R (YSQ)	*PL5*	1,584	BV097085
Ha-NT5R2	Ha-P1	R (YSQ)	*PL5*	2,021	BV097086
Ha-NT5S1	Ha-P1	S (OC)	*PL5*	1,303	BV097087
Ha-NT5S2	Ha-P1	S (OC)	*PL5*	1,424	BV097076
Ha-NT5S3	Ha-P2	S (OC)	*PL5*	387	BV097077

from potato (Ballvora et al. 2002) and rp3 from maize (Webb et al. 2002). The CC structure is followed by an NBS region that contains three ATP/GTP binding motifs known as the kinase-1a or phosphate-binding loop (P-loop); kinase-2 and kinase-3a motifs. The sequence GVGKTT of the Ha-NTIR11g and Ha-NTIR3A matches the generalized consensus (GVGKTT) (Meyers et al. 1999) for the kinase 1a P-loop. This is followed by a kinase 2 domain (LLVLDDVW) and the kinase-3a domain (GSRIIITTRD). The C terminus of the both these two RGAs contains leucine rich repeats, although most of them are imperfect.

A previous work on the major locus (*Pl6*) on linkage group 1 showed that it may contain at least 11 tightly linked genes each giving resistance to a different downy mildew race. Cloning and sequencing of 13 STSs within

this locus indicated that it contains conserved genes belonging to the TIR-NBS-LRR class of plant resistance genes (Bouzidi et al. 2002). The *Pl5*/*Pl8* locus is a second major locus for resistance to downy mildew in sunflower. The amino acid sequence analysis and homology comparison with other resistance genes indicated that this locus belongs to a different class of resistance genes (CC-NBS-LRR), which confirms the finding of Radwan et al. (2003). The present results and those of Bouzidi et al. (2002) indicate that in sunflower there are at least two regions controlling resistance to the same races of *P. halstedii* and these regions may contain different types of NBS-LRR sequences. This raises the interesting question how different types of NBS-LRR confer resistance against the same races. There are differences between the TIR-NBS-LRR sequence (*Pl6* locus) and non-TIR-NBS-LRR (*Pl5*/*Pl8* locus) not only in the N-terminus and the NBS but also in the LRR motifs. Whether these differences account for the recognition of different avirulence factors is questionable.

The *Pl5*/*Pl8* locus includes highly clustered genes

In this study, the 14 STS markers were all mapped to the *Pl5*/*Pl*8 locus on linkage group 6 of the RFLP map described by Gentzbittel et al. (1999), suggesting that this locus contains highly clustered genes. These STS markers were located within genetic distances of about 13.6 cM and 16.7 cM for the *Pl5* and *Pl8* loci, respectively. This large genetic distance may suggest that the *Pl5*/*Pl8* locus exhibits a high degree of recombination and/or that it is very large and complex. Classical genetic and molecular data show that genes determining disease resistance in plants are frequently clustered in the genome (Michelmore and Meyers 1998). For example, the *Dm3* downy mildew resistance locus of lettuce contains 32 NBS-LRR encoding genes and is spread over several megabases of one chromosome (Meyers et al. 1998; Shen et al. 2002). In sunflower, a major cluster for resistance to downy mildew has been described on linkage group 1 (Bouzidi et al. 2002), with genes also conferring resistance to all known races of downy mildew. In addition, Yu et al. (2003) mapped SCAR markers that had been found to be linked to the rust resistance genes *R1* and *Radv* (Lawson et al. 1998) on linkage groups 8 and 13 of the Yu et al. (2003) map that correspond respectively to linkage groups 1 and 6 of the Gentzbittel et al. (1999) map. As stated by these authors, the two linkage groups contain resistance both to *P. halstedii* and *Puccinia helianthi*.

Combining these two loci in marker-assisted selection programs

Numerous molecular markers closely linked to resistance genes have been recently developed in many plants; 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 (Heyes and Saghai Maroof, 2000) and the *Rph7.g* locus in barley (Brunner et al. 2000). In sunflower, 13 STSs markers were located within a genetic distance of about 3 cM of the *Pl6* locus on linkage group 1 (Bouzidi et al. 2002). Gedil et al. (2001) then Radwan et al. (2003) have described markers linked to the *Pl5/Pl8* locus. However, detection of these markers requires radiolabeled probes, which preclude their use in a vast marker-assisted selection program involving thousands of individuals. In the present study, RGA full-length sequences have been exploited to develop specific markers for the *Pl5/Pl8* locus. Because only a few primer combinations have been tested, thus the sequences provided here could be used to develop more primer pairs as necessary. The availability of six primer pairs and 14 specific PCR-based markers for the *Pl5*/*Pl8* locus should facilitate the selection and introgression of these resistance genes into new varieties. Moreover, the *Pl5/Pl8* specific primers developed here and those developed by Bouzidi et al. (2002) for the *Pl6* locus share similar characteristics, such as primer annealing temperatures and total PCR cycle numbers, which makes them compatible for multiplexing and automated PCR to introduce two different classes of resistance gene analogs in the same elite sunflower variety.

Acknowledgements The first author thanks the Egyptian Ministry of Higher Education for a Doctoral Scholarship. We thank F. Vear and D. Tourvieille de Labrouhe for critical reading of the manuscript and PROMOSOL for financial support.

References

Altschul SF, Madden TL, Schaffer AA, 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

Aravind L, Dixit VM, Koonin EV (1999) The domains of death: evolution of the apoptosis machinery. Trends Biochem Sci 24:47–53

Ballvora A, Ercolano MR, Welis J, Meksem K, Bormann CA, Oberbagemann P, Salamini F, Gebhardt C (2002) The *R1* gene for potato resistance to late blight (*Phytophthora infestans*) belongs to the leucine zipper/NBS/LRR class of plant resistance genes. Plant J 30:361–371

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

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Girauda J, Leung J, Staskawicz BJ (1994) *RPS*2 of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265:1856–1860

Bert PF, Tourvielle De Labrouhe D, Philippon J, Mouzeyar S, Jouan I, Nicolas P, Vear F (2001) Identification of a second linkage group carrying genes controlling resistance to downy mildew (*Plasmopara halstedii*) in sunflower (*Helianthus annuus* L.). Theor Appl Genet 103:992–997

Biezen EA van der, Jones JDG (1998) The NB-ARC domain: a novel signaling motif shared by plant resistance gene products and regulators of cell death in animals. Curr Biol 8:226–227

Bittner-Eddy P, Crute IR, Holub EB, Beynon JL (2000) *RPP13* is a simple locus in *Arabidopsis thaliana* for alleles that specify downy mildew resistance to different avirulence determinants in *Peronospora parasitica*. Plant J 21:177–188

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

Bouzidi MF, Badaoui S, Cambon F, Vear F, Tourvielle de Labrouhe D, Nicolas P, Mouzeyar S (2002) Molecular analysis of a major locus for resistance to downy mildew in sunflower with specific PCR-based markers. Theor Appl Genet 104:592–600

Brunner S, Keller B, Feuillet C (2000) Molecular mapping of the *Rph7.g* leaf rust resistance gene in barley (*Hordeum vulgare* L.). Theor Appl Genet 101:783–788

Flor HH (1955) Host-parasite interaction in flax rust: its genetics and other implications. Phytopathology 45:680–685

Gedil MA, Slabaugh MB, Berry S, Johnson R, Michelmore R, Miller J, Gulya T, Knapp SJ (2001) Candidate disease resistance genes in sunflower cloned using conserved nucleotide binding site motifs: genetic mapping and linkage to the downy mildew resistance gene *Pl1*. Genome 44:205–212

Gentzbittel L, Mestries E, Mouzeyar S, Mazeyrat F, Badaui 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

Grant MR, Godiard L, Straube E, Ashfield T, Lewald J, Sattler A, Innes RW, Dangl JL (1995) Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. Science 269:843–846

Hammond-Kosack KE, Jones JDG (1997) Plant disease resistance genes. Annu Rev Plant Physiol Mol Biol 48:575–607

Hammond-Kosack KE, Parker JE (2003) Deciphering plant-pathogen communication: fresh perspectives for molecular resistance breeding. Curr Opin Biotechnol 14:177–193

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

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

Jones DA, Thomas CM, Hammond-Kosack KE, Balintkurti PJ, Jones JDG (1994) Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. Science 266:789–793

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 *RPS*2 and the tobacco viral resistance gene *N*. Plant Cell 7:1195–1206

Lawson WR, Goulter KC, Henry RJ, Kong GA, Kochman JK (1998) Marker-assisted selection for two rust resistance genes in sunflower. Mol Breed 4:227–234

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

Meyers BC, Dickerman AW, Michelmore RW, Pecherer RM, Sivaramakrishnan S, Sobral B, Young ND (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide binding super family. Plant J 20:317–332

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

Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease resistance genes by bulked segregant analysis: a rapid method to markers in specific genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

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

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

Mindrinos M, Katagiri, Yu GL, Ausubel FM (1994) The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide binding site and leucine-reach repeats. Cell 78:1089–1099

Mouzeyar S, Tourvieille De Labrouhe D, Vear F (1994) Effect of host-race combination on resistance of sunflower (*Helianthus annuus* L.) to downy mildew (*Plasmopara halistedii*). J Phytopathol 141:249–258

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

Ori N, Eshed Y, Paran I, Presting G, Aviv D, Tanksley S, Zamir D, Fluhr R (1997) The *I2C* family from the wilt disease resistance locus *I2* belongs to the nucleotide binding, leuciene-rich repeat superfamily of plant resistance genes. Plant Cell 9:521–532

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

Parker JE, Holub EB, Frost LN, Falk A, Gunn ND, Daniels MJ (1996) Characterization of eds 1, a mutation in *Arabidopsis* suppressing resistance to *Peronospora parasitica* specified by several different *RPP* genes. Plant Cell 8:2033–2046

Radwan O, Bouzidi MF, Vear F, Philippon J, Tourvieille de Labrouhe D, Nicolas P, Mouzeyar S (2003) Identification of non-TIR-NBS-LRR markers linked to *PL5/PL8* locus for resistance to downy mildew in sunflower. Theor Appl Genet 106:1438–1446

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 Sci USA 81:8014–8018

Sato S, Nakmura Y, Kaneko T, Katoh T, Asamizu E, Tabata S (2000) Structural analysis of *Arabidopsis thaliana* chromosome 3 1. Sequence features of the regions of 4,504,864 bp covered by 60 P1 and TAC clones. DNA Res 7:131–135

Sela-Burrlage MB, Budai-Hadrian O, Pan Q, Carmel-Goren L, Vunsch R, Zamir D, Fluhr R (2001) Genome wide dissection of *Fusarium* resistance in tomato reveals multiple complex loci. Mol Genet Genomics 265:1104–1111

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

Shen KA, Chin DB, Arroyo-Garcia R, Ochoa OF, Lavelle DO, Wroblewski T, Meyers BG, Michelmore RW (2002) *Dm3* is one member of a large constitutively expressed family of nucleotide binding site-leucine-rich repeat encoding genes. Mol Plant Microbe Interact 3:251–261

Song WY, Xang GL, Chen LL (1995) A receptor kinase-like protein encoded by the rice resistance gene, *Xa21*. Science 270:1804–1806

Staskawicz BJ, Ausubel FM, Baker BJ, Ellis JG, Jones JDG (1995) Molecular-genetics of plant disease resistance. Science 268:661–667

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*). Theor Appl Genet 95:584–589

Webb CA, Richter TE, Collins NC, Nicolas M, Trick HN, Pryor T, Hulbert SH (2002) Genetic and molecular characterization of the maize *rp3* rust resistance locus. Genetics 162:381–394

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 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 rice *Xa21* disease resistance gene. Theor Appl Genet 93:1119–1122

Yu JK, Tang S, Slabaugh MB, Heesacker A, Cole G, Herring MJ, Soper J, Han F, Chu WC, Webb DM, Thompson L, Edwards KJ, Berry S, Leon A, Olungu C, Maes N, Knapp SJ (2003) Towards a saturated molecular genetic linkage map for cultivated sunflower. Crop Sci 43:367–387