Theor Appl Genet (2001) 102:555–563

J.G. Vicente · G.J. King

Characterisation of disease resistance gene-like sequences in *Brassica oleracea* L.

Received: 12 November 1999 / Accepted: 19 June 2000

Abstract Several cloned disease resistance genes from a wide range of plant species are known to share conserved regions with similar structural motifs. Degenerate primers based on conserved sequences of the nucleotide binding site of the genes *RPS2*, *N* and *L6* were used for polymerase chain reaction (PCR) amplification from genomic DNA of two doubled haploid lines of *Brassica oleracea*. Sequences of amplified products were highly variable, but most of them showed similarity to known disease resistance genes, including *RPS5*, *RPS2* and *N*, and to disease resistance gene-like sequences (RGLs) from different species. Primers based on *B. oleracea* sequences amplified five groups of RGLs. Products were mapped through cleaved amplified polymorphic sequence assays onto four different linkage groups of *B. oleracea*. PCR amplification from cDNA and allele analysis indicated that four locus-specific RGL fragments are expressed in cauliflower. Screening of a *B. oleracea* bacterial artificial chromosome library (BAC) with four *B. oleracea* RGL probes identified a small number of clones, suggesting that the four RGLs may not be highly copied. Screening of a BAC library of *A. thaliana* with the same probes identified clones that mapped onto four different chromosomes. These map positions correspond to known disease resistance loci of *A. thaliana*.

Keywords *Brassica* · Candidate genes · Genetic mapping · Nucleotide binding site-leucine rich repeat (NBS-LRR) · Degenerate oligonucleotides

Communicated by J.W. Snape

The sequence data reported will appear in EMBL GenBank and DDBJ databases under the accession numbers AJ250321 to AJ250326

J.G. Vicente (✉) · G.J. King
Horticulture Research International, Wellesbourne,
Warwickshire CV35 9EF, UK
e-mail: joana.vicente@hri.ac.uk

Introduction

In recent years, several genes involved in plant disease resistance have been cloned from a wide range of plant species. The R genes were grouped into different classes based on structural similarities of their predicted protein products (reviewed by Staskawicz et al. 1995; Hammond-Kosack and Jones 1997). The largest class includes gene products that contain a putative nucleotide binding site (NBS) and leucine-rich repeats (LRR) of various lengths. Genes from the NBS-LRR class are widely distributed in dicotyledon and monocotyledon plant species and confer resistance to a variety of plant pathogens including bacteria, fungi, viruses, nematodes and insects. Some of these genes, such as *N* (Whitham et al. 1994) and *L6* (Lawrence et al. 1995), have N-terminal homology to the Toll protein of *Drosophila* and the interleukin-1 receptor of mammals (TIR), while others, such as *RPS2* (Bent et al. 1994; Mindrinos et al. 1994) and *RPS5* (Warren et al. 1998), contain leucine zipper (LZ) motifs in the N-terminal region. Meyers et al. (1999) described a number of motifs present on the NBS region of TIR and non-TIR genes.

Due to the low overall sequence similarity observed between R genes, heterologous hybridisation, using cloned genes as probes, is unlikely to be a successful strategy for identifying R gene candidates (Leister et al. 1996; Shen et al. 1998). Polymerase chain reaction (PCR) amplification using degenerate primers based on conserved motifs of NBS-LRR genes is a more sensitive and efficient method to identify and characterise resistance gene-like sequences (RGLs) (Yu et al. 1996). This approach has been used successfully in the monocotyledon species maize, wheat, barley and rice (Collins et al. 1998; Seah et al. 1998; Leister et al. 1999; Mago et al. 1999) and in dicotyledon species such as potato (Leister et al. 1996), soybean (Kanazin et al. 1996; Yu et al. 1996), lettuce (Shen et al. 1998), *Arabidopsis thaliana* (Botella et al. 1997; Aarts et al. 1998; Speulman et al. 1998) and *Brassica napus* (Joyeux et al. 1999). The primers used in most of these studies were based on mo-

Table 1 Oligonucleotide primers used in this study

Primer pair	Sequences 5´3´	Based on	Reference
Degenerate primers			
I	F – ATA TGC GGC CGC GGT GGG GTW GGK AAR ACN AC R – TAT AGC GGC CGC IAR IGC IAR IGG IAR NCC	P-loop and GLPL motif of *N* and *RPS2*	Speulman et al. (1998)
II	F – GGI GGI GTI GGI AAI ACI AC R – ARI GCT ARI GGI ARA CC	P-loop and GLPL motif of *N*, *RPS2* and *L6*	Kanazin et al. (1996)
Primers based on *B. oleracea* RGLs			
III	F – AAR ACN ACT CTA CTC ACA CAA R – AGC AAC TAT GCY TGC AAG C	Sequences of 5 cloned fragments obtained with primer pair I (e.g. BoRGL-Ia)	This study
IV	F – ARA CYA CTC TTC TTG CGC AG R – CAG CGA CTT GTT TTG CGA GA	Sequences of 3 cloned fragments obtained with primer pair I (e.g. BoRGL-Ib)	This study
V	F – CGA CGR TGC TGA CTC AGG TTA R – CAC TTG GCT GCG ACC TCC TG	Sequences of 15 cloned fragments obtained with primer pair II (e.g. BoRGL-IIa)	This study
VI	F – CGA CAC TTT TCA AGA AAR TAC R – GAC ATT YTT CAG CAA CCT TTC	Sequences of 2 cloned fragments obtained with primer pair II (e.g. BoRGL-IIb)	This study

tifs of the NBS domain, including the P-loop (kinase-1a) and the GLPL motif (Meyers et al. 1999). In several studies, mapped RGLs showed close linkage to disease resistance loci indicating that this approach can contribute to the identification of candidate R genes in different species and can provide alternative starting points for cloning strategies.

Brassica oleracea is a diploid species which includes many important vegetable crops such as cabbage, broccoli, cauliflower, Brussels sprouts and kale. To date a limited number of sources of resistance to pathogens have been identified in *B. oleracea*, and only a few quantitative trait loci (QTL) that control resistance to black rot and clubroot have been mapped (Camargo et al. 1995; Grandclément and Thomas 1996; Voorrips et al. 1997; Moriguchi et al. 1999). *Brassica* species are closely related to the model plant *Arabidopsis thaliana* from which a large number of R genes have been mapped and cloned (Holub 1997). There are estimated to be 200 NBS-encoding genes in *A. thaliana* (Meyers et al. 1999). There is also evidence of a relatively high homology between coding sequences of *A. thaliana* and *Brassica*. In addition, there is extensive collinearity between the two genomes over regions covering as much as 30 cM in *A. thaliana* (Cavell et al. 1998), and a high proportion of loci in *A. thaliana* seem to be present in at least three copies in the *Brassica* genome (Kowalski et al. 1994; Lagercrantz 1998).

In the study reported here our objective was to develop a strategy based upon sequence homology and genome collinearity between *B. oleracea* and *A. thaliana* in order to obtain a preliminary characterisation of the pool of R gene sequences present in the *B. oleracea* genome. We applied a PCR approach using heterologous degenerate primers, based on conserved motifs of TIR and non-TIR R genes, to isolate and sequence RGLs of *B. oleracea*. To obtain a detailed characterisation we used this preliminary information to design more specific *B. oleracea* primers, which enabled us to clone, sequence and map locus-specific fragments using cleaved amplified polymorphic sequence (CAPS) assays. In addition, we were able to confirm whether these specific sequences were expressed. Finally, we determined the corresponding map position of orthologous sequences in *A. thaliana*.

Materials and methods

Plant material and genomic DNA extraction

All genomic DNA was extracted from fresh or freeze-dried leaf material using a cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle 1990) with minor modifications. DNA of the parental doubled haploid (DH) lines A12DHd (*B. oleracea* var. *alboglabra*) and GDDH33 (*B. oleracea* var. *italica*) was used as template for PCR. A subset of 120 DH lines of a mapping population derived from a cross between A12DHd and GDDH33 was used for mapping studies (Bohuon et al. 1996; Sebastian et al. 2000). Two mapping assays were repeated using 97 DH lines derived from a cross between the DH line CA25 from Nedcha (*B. oleracea* var. *botrytis*) and the DH line AC498 from Gower (*B. oleracea* var. *gemmifera*) (Sebastian et al. 2000).

PCR analysis

Primer pairs are presented in Table 1. Two pairs of degenerate PCR primers (I and II) designed by Speulman et al. (1998) and Kanazin et al. (1996), based on the P-loop and the GLPL motif of *RPS2*, *N* and *L6*, were used for amplifying related sequences from *B. oleracea* A12DHd and GDDH33. The initial PCR was in a total volume of 20 µl containing 100 ng genomic DNA, 20 pmol each primer (Genosys), 1.5 m*M* $MgCl_2$, 400 µ*M* dNTPs and 0.6 U *Taq* DNA polymerase (Gibco BRL), overlaid with mineral oil. PCR amplification was performed in an Hybaid Omni-Gene thermocycler. Standard cycling conditions were 3 min at 94°C, followed by 35 cycles of a 1-min annealing, 1 min at 72°C and 1 min at 94°C and a final cycle of a 2-min annealing and 5 min at 72°C. A PCR Optimizer Kit (Invitrogen) was used to optimise the pH and magnesium concentration. Different annealing temperatures and concentrations of genomic DNA and primers were also tested. Products were separated on 1.5% SeaKem agarose gels with ethidium bromide at 0.5 µg/ml in 1×TAE and visualised under UV light. A

Table 2 Conditions of PCR, representative RGL fragments cloned and sequenced and restriction enzymes used to define marker loci in CAPS mapping assays

Primer pair	Annealing temperature (°C)	Optimised PCR conditions[a]			Representatives of cloned and sequenced RGL fragments	Restriction enzyme for mapping assays
		DNA (ng)	Primers (pmol)	$MgCl_2$ (nmol)		
I	52	100	20	24	*BoRGL-Ia, BoRGL-Ib, BoRGL-Ic, BoRGL-Id*	–
II	55	75	8	50	*BoRGL-IIa, BoRGL-IIb*	–
III	60	100	8	40	*BoRGL-IIIa (similar to BoRGL-Ia)*	*Sau*3 A
IV	62	100	8	40	*BoRGL-IVa (similar to BoRGL-Ib)*	*Nla*IV
V	65	100	8	50	*BoRGL-Va (similar to BoRGL-IIa)*	*Msp*I
VI	56	100	8	40	*BoRGL-VIa (similar to BoRGL-IIb) and BoRGL-VIb (longer fragment)*	*Hae*III

[a] In a total volume of 20 µl including 400 µ*M* dNTPs and 0.8 U *Taq* polymerase

100-bp DNA ladder (Gibco BRL) was used to estimate the size of each amplified DNA fragment.

Four pairs of more specific primers (Table 1) were designed based on distinct groups of *B. oleracea* sequences that showed similarity with non-TIR R genes (e.g. *RPS5* and *RPS2*). Primer design was based upon sequences nested within, and adjacent to, the P-loop and GLPL motifs.

Cloning and sequence analysis of PCR products

PCR amplification was performed using the conditions described in Table 2. PCR products were cloned into the plasmid pCR2.1 using an Original TA Cloning Kit (Invitrogen). Plasmid DNA was isolated using a Qiagen-tip 20 Plasmid Kit and used as template in sequencing reactions using the BigDye™ Terminator Cycle Sequencing Ready Reaction with AmpliTaq® DNA Polymerase kit from ABI PRISM™ (Perkin Elmer) with M13 forward or M13 reverse primers. Sequences of PCR products were determined using an ABI PRISM™ 377 DNA sequencer.

DNA sequence analysis and comparisons were carried out using DNASTAR™ software. Searches for similarity between the predicted protein products of clones and proteins on the database were carried out using BLASTX (Gish and States 1993).

Locus-specific assays using cleaved amplified polymorphic sequence (CAPS) markers

PCR products obtained with four pairs of *B. oleracea* primers (III–VI) from DH lines were digested with selected restriction enzymes (Table 2). Products were separated on 1.8% agarose gels in 0.5×TBE with ethidium bromide. Segregating bands were scored and assigned to an integrated map (Sebastian et al. 2000) using JOINMAP™ 2.0 (Stam and Van Ooijen 1995).

Expression of disease resistance gene-like (RGL) fragments in cauliflower

cDNA derived from curd tissue of the DH line CA25 of Nedcha cauliflower (*B. oleracea* var. *botrytis*) was kindly donated by Erik Kop at HRI-Wellesbourne. This was used as a template in PCR with four pairs of *B. oleracea* primers (III–VI). For comparison, genomic DNA from the same genotype was also used in PCR.

Bacterial Artificial Chromosome (BAC) – filter hybridisation and characterisation of clones

Two BAC libraries were probed. The 'BoB' BAC library of *B. oleracea* A12DHd (Ryder and King, unpublished; http://hbz.tamu.edu/bacindex4.html) contains 26,000 clones with an average insert size of 120 kb and is estimated to represent 5 genome equivalents. The IGF-BAC library of *A. thaliana* contains 10,752 clones with an average insert size of 100 kb and represents 7.5 haploid genome equivalents (Mozo et al. 1998a). Filters and specific clones from this library were obtained from the Arabidopsis Biological Resource Center (ABRC), Ohio, USA. Probes were prepared from four cloned A12DHd PCR products obtained with *B. oleracea* primer pairs III and VI and were labelled with [^{32}P] using a nick translation kit (Amersham Life Sciences). Pre-hybridisation and hybridisation were overnight at 65°C in 0.5 *M* sodium phosphate, 7% (w/v) SDS, 1% (w/v) BSA, 1 m*M* EDTA with 10 mg of salmon sperm DNA. Filters were washed either in 0.5×SSC and 0.1% SDS at 65°C for 20 min (*B. oleracea*) or at low stringency in 2×SSC and 0.1% SDS at 65°C for 5 min (*A. thaliana*) and exposed to Kodak BioMax MS-1 film between 2 and 8 days. *A. thaliana* BAC clones with clear hybridisation signals were mapped using information available at http://www.mpimp-golm.mpg.de/101/mpi_mp_map/bac.html (Mozo et al. 1998b) and the physical map available at http://genome-www3.stanford.edu/.

Results

PCR identification of *Brassica oleracea* RGLs using degenerate primers

RGLs were amplified from genomic DNA of *B. oleracea* A12DHd and GDDH33 with two pairs of degenerate primers (Table 1). Using optimised PCR conditions (Table 2) we obtained products of approximately 530 bp and 610 bp with primer pairs I and II, respectively. In total 37 PCR products from A12DHd and GDDH33 were cloned and sequenced. Only 3 sequences (obtained with primer pair I) did not show similarity with sequences related to known R genes. Figure 1 presents an alignment of deduced amino acid sequences between the P-loop and the GLPL motif of 6 *B. oleracea* RGLs (BoRGLs) representative of the diversity of the cloned fragments. These sequences were aligned with sequences of three R genes that showed the highest similarity with the BoRGLs [*RPS5* (EMBL accession AF074916), *RPS2* (U12860) and *N* (U15605)] and the deduced protein of an RGL from *A. thaliana*, pNd11 (U97223).

Sequences of cloned fragments obtained with primer pair I were highly variable. However, two groups of fragments with more than 97% nucleotide identity were identified. One group consisted of 5 fragments (from A12DHd and GDDH33) and the other of 3 fragments (from GDDH33). BoRGL-Ia and BoRGL-Ib are repre-

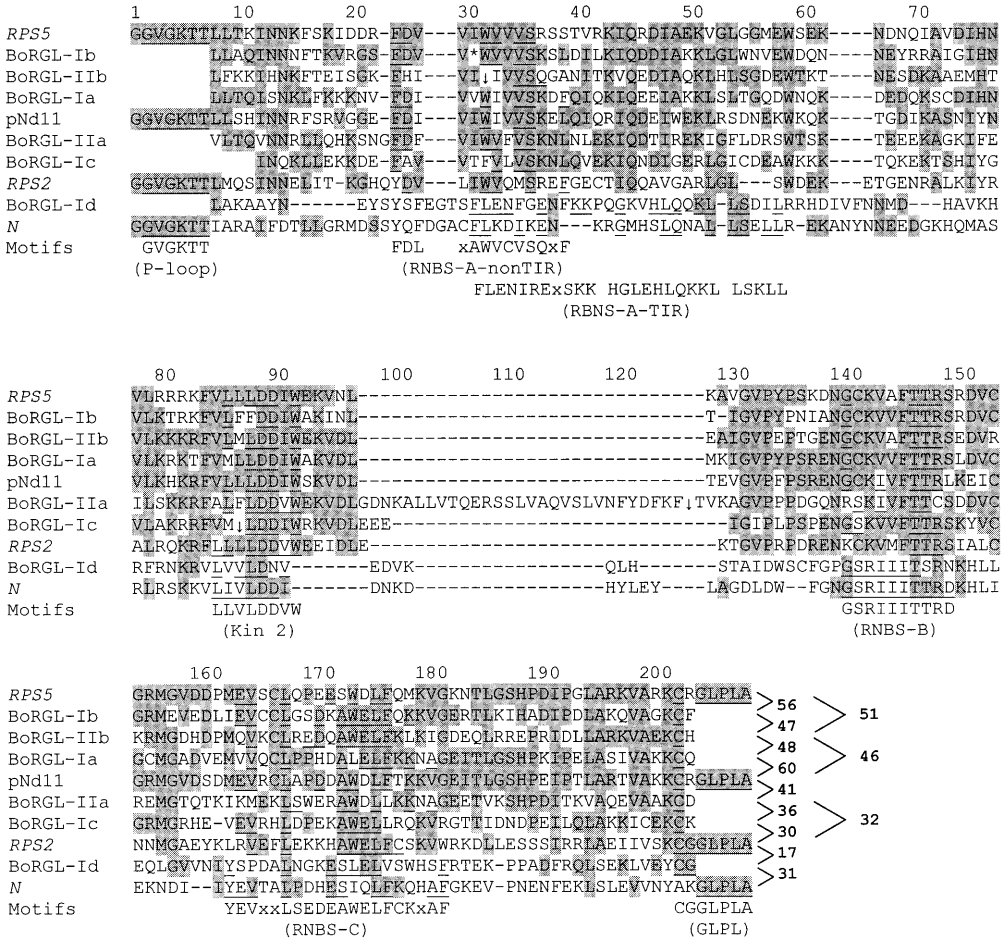

Fig. 1 Deduced amino acid sequences of *Brassica oleracea* resistance gene-like cloned PCR fragments obtained from genomic DNA with degenerate primers based on the P-loop and GLPL motifs. BoRGL-I and BoRGL-II sequences were obtained with primer pair I and II, respectively. BoRGL-Ia, BoRGL-Ic, BoRGL-IIa and BoRGL-IIb are from the parental line A12DHd. BoRGL-Ib and BoRGL-Id are from the parental line GDDH33. The sequences are aligned with similar regions of NBS-LRR resistance gene products (*RPS5*, *RPS2* and *N*) and a deduced amino acid sequence of a resistance gene-like fragment of *A. thaliana* (pNd11). The sequences were aligned by the Clustal method with PAM250 residue weight table and ordered according to their similarity. Residues that match the consensus are *shaded*. The presence of a frame shift is indicated by an *asterisk* (*), and stop codons are indicated by an *arrow* (↓). The level of amino acid identity as a percentage is indicated at the *end* of the *bottom panel*. Amino acids identical to motifs described by Meyers et al. (1999) found in these sequences are *underlined*. *B. oleracea* nucleotide and/or amino acid sequences are to be found at the EMBL, GenBank and DDBJ databases under the accession numbers AJ250321 to AJ250326

sentative of these two groups. The deduced translation products of BoRGL-Ia and BoRGL-Ib have 56% amino acid identity with *RPS5*. BoRGL-Ia has 60% amino acid identity with pNd11. The sequences from the BoRGL-Ib group have a 2-bp deletion when aligned with the sequences from the BoRGL-Ia group, resulting in a frame shift at amino acid position 31. The alignment presents the sequences of two possible reading frames, 5′ of that deletion and 3′ of that deletion, which show similarity to R genes. Other fragments such as BoRGL-Ic and BoRGL-Id are examples of different sequences obtained with primer pair I. These show less than 25% nucleotide identity between them and less than 50% identity with the other 2 sequences. BoRGL-Ic has 40% and 30% amino acid identity with *RPS5* and *RPS2*, respectively. This sequence has a stop codon at the amino acid position 87. BoRGL-Id has 31% amino acid identity with *N* and less than 18% with *RPS2* and *RPS5*. This sequence has shown 65–35% amino acid identity to several RGL products from different species, including *A. thaliana* (AF039377; Aarts et al. 1998), pea (AF123699), sunflower (U96642), soybean (U55810; Kanazin et al. 1996) and potato (U60080; Leister et al. 1996).

Sequences of cloned fragments obtained with primer pair II form only two distinct groups sharing less than 45% nucleotide identity. The largest group represented by BoRGL-IIa includes 15 fragments from A12DHd and GDDH33 with more than 75% nucleotide identity. BoRGL-IIa is longer than the other sequences and has 39% and 32% amino acid identity with *RPS5* and *RPS2*, respectively. The second group, represented by BoRGL-IIb, only includes 2 fragments from A12DHd. BoRGL-IIb has 51% amino acid identity with *RPS5*. The sequences BoRGL-IIa and BoRGL-IIb have stop codons at amino acid positions 127 and 32, respectively.

Between the P-loop and the GLPL motifs, all representative BoRGL amino acid sequences included segments that showed similarity with motifs described by Meyers et al. (1999) (Fig. 1). All BoRGLs have the motifs Kin 2, RNBS-B and RNBS-C. All, except BoRGL-Id, have a RNBS-A-nonTIR motif. BoRGL-Id has a RBNS-A-TIR motif (as in the resistance gene *N*).

Design of *Brassica oleracea* RGL primers and sequencing of products

To further characterise specific BoRGLs, we designed four pairs of primers based on groups of *B. oleracea* sequences that showed similarity with non-TIR R genes. Primer design was based upon sequences adjacent to the P-loop and the GLPL motifs (Table 1). Under optimised PCR conditions (Table 2) products of approximately 480 bp were amplified with primer pairs III and IV, products of 590 bp were amplified with primers V and products of 480 bp (from A12DHd and GDDH33) and 800 bp (from GDDH33 only) were amplified with primers VI. Twenty-two cloned fragments of A12DHd and 14 of GDDH33 were sequenced. All sequences showed similarity to known R genes (including *RPS5* and *RPS2*). Five distinct groups of fragments were identified. Sequences from each of four groups (represented by BoRGL-IIIa, BoRGL-IVa, BoRGL-Va, BoRGL-VIa) were very uniform, showing more than 80% nucleotide identity. As expected these four groups correspond to the sequences BoRGL-Ia, BoRGL-Ib, BoRGL-IIa and BoRGL-IIb obtained previously. The fifth group (e.g. BoRGL-VIb) includes 2 longer products (800 bp) obtained from GDDH33 with primer pair VI. These sequences also showed similarity to known R gene-related sequences.

Mapping of *Brassica oleracea* RGL loci

Restriction enzyme digestion of PCR products obtained with degenerate primer pairs (especially with primer pair I) produced many fragments with some of the enzymes tested (e.g. *Taq*I), indicating the existence of a mixture of PCR products even in what appeared to be a single band on the gel.

Restriction enzyme digestion of PCR products obtained with BoRGL primer pairs III–VI produced a small number of fragments, indicating that these products were more homogeneous than the products obtained with degenerate primers. Polymorphisms between A12DHd and GDDH33 PCR products were obtained with the restriction enzymes indicated in Table 2. CAPS marker assays were used to score segregation in one or two mapping populations. Following linkage analysis of the AxG population results, products obtained with primer pair III (e.g. BoRGL-IIIa) were mapped on linkage group O4, products obtained with primer pair IV (e.g. BoRGL-IVa) were mapped on linkage group O8, products obtained with primer pair V (e.g. BoRGL-Va) were mapped on linkage group O1 and products obtained with primer pair VI (e.g. BoRGL-VIa) were mapped on linkage group O9 (Fig. 2). Polymorphism between the N×G parent DH lines was found in products obtained with primer pairs IV and V (Fig. 3). The locus positions on linkage groups O1 and O8 fell within identical mapping intervals for both the N×G population and the A×G population.

Expression of *Brassica oleracea* RGLs in cauliflower

Using the same four pairs of locus-specific *B. oleracea* primers (III–VI), we obtained amplified products of the expected size, equivalent to products obtained with A12DHd and GDDH33, from both genomic DNA and cDNA of the cauliflower DH line CA25. These products were compared after digestion with the restriction enzymes used in the mapping assays (Table 2). The digested products amplified from both cDNA and genomic DNA with primer pairs IV and V showed the same number of fragments (Fig. 3). The digested products amplified from cDNA with primer pairs III and VI had fewer fragments

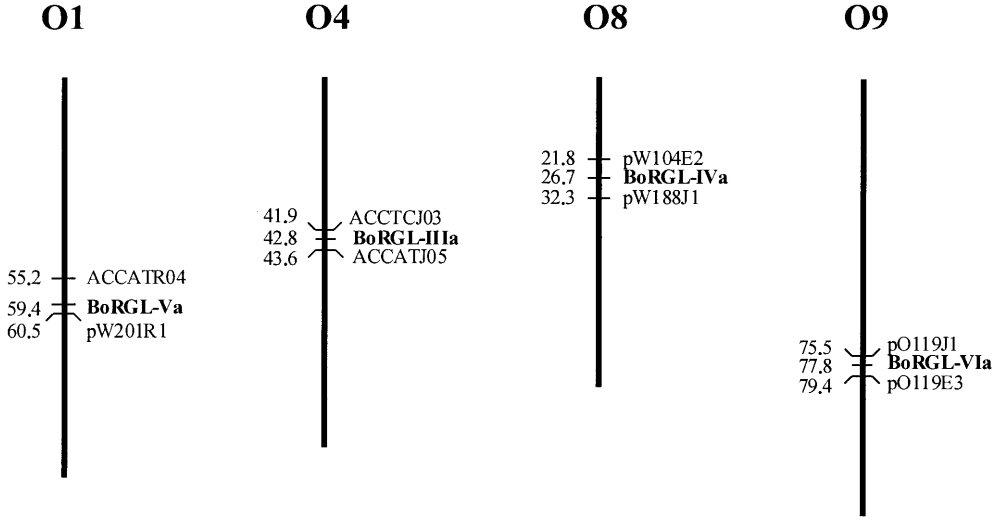

Fig. 2 Map position of *Brassica oleracea* RGL loci in four of the nine linkage groups. *B. oleracea* linkage groups O1, O4, O8 and O9 are represented by *vertical lines*. On the *right* of each linkage group, markers identified in this study are shown in *bold* together with closely linked molecular markers; *numbers* on the *left* are genetic distances in centiMorgans (cM). Linkage groups correspond to those described in Bohuon et al. (1996) and Sebastian et al. (2000)

than products from genomic DNA. This observation suggests that two different products were amplified from the cauliflower genomic DNA and that only one of those products was expressed at detectable levels in the curd tissue sampled. The sequences of products obtained with primer pair V from both genomic and cDNA were identical to BoRGL-IIa. These results indicate that each BoRGL primer pair amplify at least 1 RGL fragment that is expressed in cauliflower curd tissue.

Identification of *Brassica oleracea* genomic clones with RGLs

In order to validate the locus specificity of the sequences obtained, we hybridised cloned probes representing the four locus-specific BoRGLs against the BoB BAC library of *B. oleracea* A12DHd. While between 3 and 38 BAC clones were detected (Table 3), only 17 (19%) of these clones showed a strong signal. Twelve BAC clones identified with the BoRGL-IIIa probe and 3 BAC clones identified with the BoRGL-Va probe were selected to establish whether they contained fragments which corresponded with those previously amplified from genomic DNA, and then used in the locus-specific CAPS assays. This was confirmed by restriction enzyme digest patterns and by sequencing. PCR products of the expected size were amplified from only 2 BAC clones with primer pair III (BoB11O13 and BoB60EO3) and 1 BAC clone with primer pair V (BoB13D19). The *Sau3A* digested products from BoB11O13 showed the A12DHd genomic pattern; in contrast, the digested products from BoB60EO3 showed fewer fragments than A12DHd genomic DNA. The sequences of the PCR products from BoB11O13 and BoB60EO3 showed 99% and 93% nucleotide identity with BoRGL-Ia. The products of BoB13DO9 digested with *Msp*I had the same fragment sizes as A12DHd. The sequence of the PCR product from BoB13DO9 is identical to BoRGL-IIa. These results indicate that BoB11O13 and BoB13DO9 contain the loci mapped on linkage group O4 and O1, respectively.

PCR products of the expected sizes were obtained with the degenerate primer pair I from all the selected clones that hybridised with the BoRGL-IIIa probe, with the exception of BoB34I 7; PCR products were obtained with degenerate primer pair II only from BoB13D19. These results indicate that most of the clones that hybridised with the probes contain RGLs but that only a small number of these clones contain the specific fragments that have been sequenced and mapped.

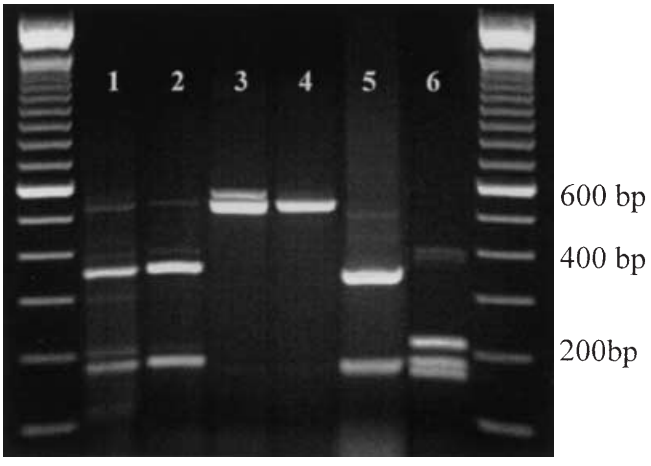

Fig. 3 PCR products amplified with primer pair V digested with *Msp*I. The template used was: *1* cDNA from CA25, *2* genomic DNA from CA25, *3* genomic DNA from line AC498, *4* genomic DNA from GDDH33, *5* genomic DNA from A12DHd, *6* DNA from the IGF-BAC of *A. thaliana* F20P7. A 100-bp ladder (Gibco BRL) was used as size markers

Physical mapping of RGLs in *Arabidopsis thaliana*

Four A12DHd cloned probes representing four locus-specific BoRGLs were used to screen the IGF-BAC library of *A. thaliana* (Mozo et al. 1998a,b). The map location of the clones that hybridised with each probe is indicated in Fig. 4. The BoRGL-IIIa and BoRGL-IVa probes hybridised with 14 IGF-BAC clones. Apart from 2 unmapped BAC clones (F9A11 and F10G7), these could be grouped in two contigs, one on chromosome I (8 clones including F22I8) and one on chromosome IV (4 clones including F24D14). The BoRGL-VIa probe identified IGF-BAC clones from the same two contigs, plus 7 clones from another contig on chromosome I (including F11P17) and 4 unmapped clones. The BoRGL-Va probe hybridised with 16 clones. Five clones belong to contigs on chromosome I (including F2B15 and F3A5), IV (F14E11) and V (F22F21), and a group of 5 clones (including F20P7) belong to a contig on chromosome III. The 6 remaining clones have not been mapped. Most of the unmapped clones have been reported to con-

Table 3 *Brassica oleracea* BoB BAC clones that hybridised with four probes from cloned BoRGL products obtained with four pairs of *B. oleracea* primers

Probe	BoB BAC hybridising clones	
	Total (strong hybridisation)	Selected clones
BoRGL-IIIa	26 (14)	BoB11O13, BoB24PO2, BoB19FO8, BoB25BO4, BoB27DO5, BoB33B23, BoB34IO7, BoB42O10, BoB46O17, BoB53L22, BoB60EO3, BoB68P19
BoRGL-IVa	23 (1)	–
BoRGL-Va	3 (1)	BoB13D19, BoB25LO1, BoB58I10
BoRGL-VIa	38 (1)	–

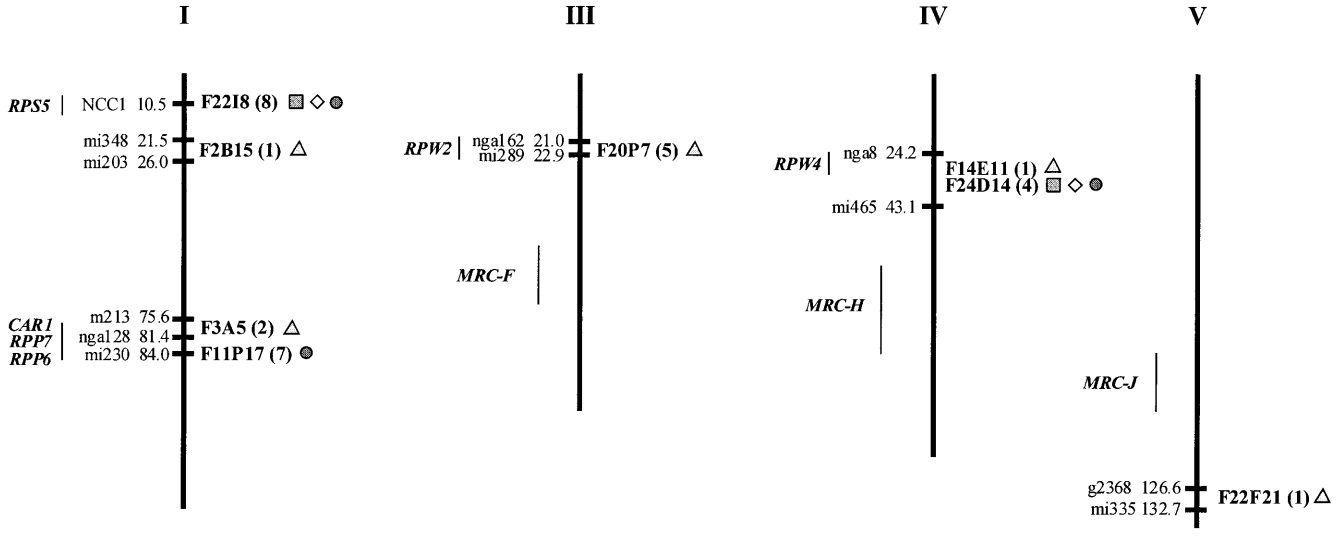

Fig. 4 Location of *Arabidopsis thaliana* IGF-BAC clones that hybridised with four *Brassica oleracea* probes from cloned PCR products: BoRGL-IIIa (▨), BoRGL-IVa (◇), BoRGL-Va (△) and BoRGL-VIa (◍). Closely linked markers are indicated according to information available at http://www.mpimp-golm.mpg.de/101/mpi_mp_map/bac.html (Mozo et al. 1998b). Mapped disease resistance genes in the same regions and major recognition complexes are also indicated (Kunkel et al. 1996; Holub 1997). The number of clones from each contig group is indicated in *brackets*. The F22I8 group also includes F11H24, F21H10, F23G11, F10F7, F27M4, F26I1 and F5O11; F3A5 also includes F22P19; the F11P17 group also includes F3D11, F17O10, F25J21, F27J4, F2M24 and F27E22; the F20P7 group also includes F19D20, F13P11, F3L4 and F5L7; the F24D14 group also includes F17M8, F22M1 and F6D3

tain nuclear repetitive sequences or 18S-25 S rDNA (Mozo et al. 1998a).

The map position of the majority of the clones could be associated with the position of functional R genes of *A. thaliana* (reviews by Kunkel 1996; Holub 1997) (Fig. 4). The clones of the F22I8 group mapped to an interval that includes the gene *RPS5*, which is involved in the recognition of *Pseudomonas syringae* pv. *phaseolicola* (Simonich and Innes 1995). The clones from the F11P17 and F3A5 groups mapped to an interval that contains *RPP6* and *RPP7*, which are involved in the recognition of *Peronospora parasitica* (Holub and Beynon 1997), and *CAR1*, which is involved in the recognition of turnip mosaic virus (Leisner et al. 1993). Additionally, the complete sequence of the clone F11P17 revealed the presence of two RGLs with similarity to *RPS2* (F11P17.9 and F11P17.10; EMBL accession AC002294). The clones of the F20P7 group mapped to an interval that includes *RPW2*, and the clones F14E11 and F24D14 mapped close to the gene *RPW4*. *RPW2* and *RPW4* are involved in the recognition of *Erisyphe cichoracearum* (Adam and Somerville 1996). To date, no functional disease resistance gene has been reported in the F22F21 region on chromosome V.

To confirm the presence of RGLs in the *A. thaliana* IGF-BAC clones that hybridised with the BoRGL probes, we used DNA from 8 IGF-BAC clones in PCR with primer pairs III (F22I8, F23G11, F5O11, F24D14) and V (F2B15, F3A5, F20P7, F5L7). Products of the expected size were obtained with primer pair V from the clones F2B15, F20P7 and F5L7. These products have an additional *Msp*I restriction site that is not present in BoRGL-IIa, BoRGL-Va and BoB13D19 (Fig. 3). The nucleotide sequences of these three *A. thaliana* products are more than 95% identical and are approximately 80% identical to BoRGL-IIa. No products were obtained with primer pair III. With the degenerate primer pair I, PCR products of the expected sizes were obtained from the IGF-BAC clones F22I8, F23G11, F5O11 and F24D14 and with degenerate primer pair II from the clone F3A5. This indicates that RGLs are present in all these clones.

Discussion

We have successfully applied a PCR strategy to isolate locus-specific RGLs from *B. oleracea* based on the use of degenerate primers designed from conserved motifs of the NBS region of known TIR and non-TIR R genes.

PCR with degenerate primers amplified a mixture of fragments, containing different sequences of approximately the same length, from *B. oleracea* A12DHd and GDDH33 genomic DNA. Several authors have reported similar results (Chen et al. 1998; Mago et al. 1999). Most of the sequences showed similarity to known R genes and at least six groups of *B. oleracea* RGLs have been identified. Many RGLs have been identified in other plant species using a similar strategy, including *A. thaliana* (Aarts et al. 1998; Speulman et al. 1998). However, in *B. napus*, only three of ten groups of sequences amplified with primers based on the LRR and NBS region of *RPS2* showed similarity with *RPS2* or RGLs from different species (Joyeux et al. 1999).

A group of cloned RGL fragments has a deletion resulting in a frame shift (BoRGL-Ib), and three others have stop codons (BoRGL-Ic, BoRGL-IIa and BoRGL-IIb). Aarts et al. (1998) also found RGLs with frame shifts and stop codons in *A. thaliana*. Evidence from the

RPP5 multigene family suggests that despite the presence of retroelement insertions and frame shift mutations, such genes may still be functional (Noël et al. 1999). Therefore, the RGLs described in this study could correspond to functional R genes of *B. oleracea* cultivars or other *Brassica* species. Alternatively, they could be closely linked to functional R genes, as functional R genes are frequently found to be replicated, clustered and closely linked to non-functional pseudogenes (Lagudah et al. 1997; McDowell 1998; Warren et al. 1998).

Several motifs typical of the NBS region of resistance genes were identified in the *B. oleracea* RGLs. Only 2 cloned fragments (e.g. BoRGL-Id) contain the RNBS-A-TIR motif. All other cloned fragments had the RNBS-A-nonTIR motif (Meyers et al. 1999), suggesting that they belong to genes (or pseudogenes) that do not contain the TIR region.

Four pairs of *B. oleracea* primers were designed based on BoRGLs that showed similarity with non-TIR R genes. CAPS assays were successfully used to map each one of the four BoRGLs to unique loci in four different linkage groups of *B. oleracea*. In contrast, using a hybridisation strategy, most of the *B. napus* restriction fragment length polymorphism (RFLP) probes used by Joyeux et al. (1999) mapped to multiple loci distributed throughout the genome.

We have not been able to associate the four BoRGL loci with the disease resistance QTL previously mapped in *B. oleracea.* However, very few functional resistances have been mapped in this species and none of these as a major gene locus. The RFLP markers identified in *B. napus* also did not co-segregate with any major gene locus that controls resistance to blackleg and white rust (Joyeux et al. 1999). Nevertheless, the BoRGLs characterised in this study are being used as markers or candidate genes in ongoing research programmes aimed at locating and cloning R genes accounting for resistance to major diseases of *Brassica*, including black rot, clubroot, downy mildew and white blister.

The amplification and restriction analysis of PCR products from cDNA of cauliflower provided evidence of the expression of four mapped BoRGL fragments. Sequencing confirmed the presence of identical sequence in cDNA and genomic DNA for one fragment. Aarts et al. (1998) also demonstrated that some *A. thaliana* RGLs are transcriptionally active, and several expressed sequences showing similarity to R genes have been mapped by Botella et al. (1997).

Four BoRGL probes hybridised with a limited number of BoB BAC clones. We only found 2 BAC clones that contain the mapped restriction sites, suggesting that they correspond to BoRGL-IIIa and BoRGL-Va. Although we expect several hundreds of NBS-encoding genes to be present in the *B. oleracea* genome, these results indicate that the specific RGLs characterised in this study are not highly replicated.

The BoRGL probes hybridised to *A. thaliana* IGF-BAC clones mapped to eight loci on four chromosomes of *A. thaliana*. Most of these loci appear to be closely linked to functional R genes and to RGLs previously mapped by Botella et al. (1997), Aarts et al. (1998) and Speulman et al (1998). The distribution of the loci detected with the non-TIR BoRGL probes is consistent with that described by Meyers et al. (1999).

Three of the BoRGL probes identified the same two contigs of overlapping IGF-BAC clones. This result is in agreement with results reported in other studies and indicates that the *Brassica* genomes are replicated compared to the *A. thaliana* genome (Kowalski et al. 1994; Lagercrantz 1998). Nevertheless, each BoRGL probe hybridised with clones from more than one region of *A. thaliana*, indicating the presence of similar RGLs in different parts of the genome.

The characterisation of RGLs from *B. oleracea* may provide additional markers to regions related with functional or non-functional R genes in this species, or in related species including *A. thaliana*. Furthermore, the analysis of related sequences from different species should provide additional information about the evolution of NBS-LRR gene families and basis of allelic variation in the *Brassica* gene pool that is of importance for crop genetics and improvement.

Acknowledgements We thank Elly Speulman and Jim Beynon for providing primer sequences prior to publication, Lee Smith and Graham Teakle for assistance with analysis of the mapping results, Carol Ryder for assistance with BAC library hybridisation work, Erik Kop for generating the cDNA and Jim Beynon for assistance with sequence analysis and review of the manuscript. This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC). Joana Vicente was supported by a grant from PRAXIS XXI (BD/3780/94).

References

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

Adam L, Somerville SC (1996) Genetic characterization of five powdery mildew disease resistance loci in *Arabidopsis thaliana*. Plant J 9:341–356

Bent AF, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ (1994) *RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes. Science 265:1856–1860

Bohuon EJR, Keith DJ, Parkin IAP, Sharpe AG, Lydiate DJ (1996) Alignment of the conserved C genomes of *Brassica oleracea* and *Brassica napus*. Theor Appl Genet 93:833–839

Botella MA, Coleman MJ, Hughes DE, Nishimura MT, Jones JDG, Somerville C (1997) Map positions of 47 *Arabidopsis* sequences with sequence similarity to disease resistance genes. Plant J 12:1197–1211

Camargo LEA, Williams PH, Osborn TC (1995) Mapping of quantitative trait loci controlling resistance of *Brassica oleracea* to *Xanthomonas campestris* pv. *campestris* in the field and greenhouse. Phytopathology 85:1296–1300

Cavell AC, Lydiate DJ, Parkin, IAP, Dean C, Trick M (1998) Collinearity between a 30-centimorgan segment of *Arabidopsis thaliana* chromosome 4 and duplicated regions within the *Brassica napus* genome. Genome 41:62–69

Chen XM, Line RF, Leung H (1998) Genome scanning for resistance-gene analogs in rice, barley, and wheat by high-resolution electrophoresis. Theor Appl Genet 97:345–355

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

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Gish W, States DJ (1993) Identification of protein coding regions by database similarity search. Nat Genet 3:266–272

Grandclément C, Thomas G (1996) Detection and analysis of QTLs based on RAPD markers for polygenic resistance to *Plasmodiophora brassicae* Woron in *Brassica oleracea* L.. Theor Appl Genet 93:86–90

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

Holub EB (1997) Organization of resistance genes in *Arabidopsis*. In: Crute IR, Holub EB, Burdon JJ (eds) The gene-for-gene relationship in plant-parasite interactions. CAB International, Wallingford, UK, pp 5–26

Holub EB, Beynon JL (1997) Symbiology of mouse-ear cress (*Arabidopsis thaliana*) and oomycetes. Adv Bot Res 24:227–273

Joyeux A, Fortin MG, Mayerhofer R, Good AG (1999) Genetic mapping of plant disease resistance gene homologues using a minimal *Brassica napus* L. population. Genome 42:735–743

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

Kowalski SP, Lan TH, Feldmann KA, Paterson AH (1994) Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveal islands of conserved organization. Genetics 138:499–510

Kunkel B (1996) A useful weed put to work: genetic analysis of disease resistance in *Arabidopsis thaliana*. Trends Genet 12:63–69

Lagercrantz U (1998) Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. Genetics 150:1217–1228

Lagudah ES, Moullet O, Appels R (1997) Map-based cloning of a gene sequence encoding a nucleotide-binding domain and a leucine-rich region at the *Cre3* nematode resistance locus of wheat. Genome 40:659–665

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

Leisner SM, Turgeon R, Howell SH (1993) Effects of host plant development and genetic determinants on the long-distance movement of cauliflower mosaic virus in Arabidopsis. Plant Cell 5:191–202

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

Leister D, Kurth J, Laurie DA, Yano M, Sasaki K, Graner A, Schulze-Lefert P (1999) RFLP and physical mapping of resistance gene homologues in rice (*O. sativa*) and barley (*H. vulgare*). Theor Appl Genet 98:509–520

Mago R, Nair S, Mohan M (1999) Resistance gene analogues from rice: cloning, sequencing and mapping. Theor Appl Genet 99:50–57

McDowell JM, Dhandaydham M, Long TA, Aarts MGM, Goff S, Holub E, Dangl JL (1998) Intragenic recombination and diversifiying selection contribute to the evolution of downy mildew resistance at the *RPP8* locus of Arabidopsis. Plant Cell 10:1861–1874

Meyers BC, Dickerman AW, Michelmore RW, Sivaramakrishnab 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

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

Moriguchi K, Kimizuka-Takagi C, Ishii K, Nomura K (1999) A genetic map based on RAPD, RFLP, isozyme, morphological markers and QTL analysis for clubroot resistance in *Brassica oleracea*. Breed Sci 49:257–265

Mozo T, Fisher S, Shizuya H, Altmann T (1998a) Construction and characterization of the IGF Arabidopsis BAC library. Mol Gen Genet 258:562–570

Mozo T, Fisher S, Meier-Ewert S, Lehrach H, Altmann T (1998b) Use of the IGF BAC library for physical mapping of the *Arabidopsis thaliana* genome. Plant J 16:377–384

Noël L, Moores TL, van der Biezen EA, Parniske M, Daniels MJ, Parker JE, Jones JDG (1999) Pronounced intraspecific haplotype divergence at the *RPP5* complex disease resistance locus of Arabidopsis. Plant Cell 11:2099–2111

Seah S, Sivasithamparam K, Karakousis A, Lagudah ES (1998) Cloning and characterisation of a family of disease resistance gene analogs from wheat and barley. Theor Appl Genet 97:937–945

Sebastian RL, Howell EC, King, GJ, Marshall DF, Kearsey MJ (2000) An integrated AFLP and RFLP *Brassica oleracea* linkage map from two morphologically distinct doubled haploid mapping populations. Theor Appl Genet 100:75–81

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

Simonich MT, Innes RW (1995) A disease resistance gene in *Arabidopsis* with specificity for the *avrPph3* gene of *Pseudomonas syringae* pv. *phaseolicola*. Mol Plant-Microbe Interact 8:637–640

Speulman E, Bouchez D, Holub EB, Beynon JL (1998) Disease resistance gene homologs correlate with disease resistance loci of *Arabidopsis thaliana*. Plant J 14:467–474

Stam P, Van Ooijen JW (1995) JoinMap™ 2.0: software for the calculation of genetic linkage maps. CPRO-DLO, Wageningen, the Netherlands

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

Voorrips RE, Jongerius MC, Kanne HJ (1997) Mapping of two genes for resistance to clubroot (*Plasmodiophora brassicae*) in a population of doubled haploid lines of *Brassica oleracea* by means of RFLP and AFLP markers. Theor Appl Genet 94:75–82

Warren RF, Henk A, Mowery P, Holub E, Innes RW (1998) A mutation within the leucine-rich repeat domain of the Arabidopsis disease resistance gene *RPS5* partially suppresses multiple bacterial and downy mildew resistance genes. Plant Cell 10:1439–1452

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

Yu YG, Buss GR, Maroof MAS (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