Theor Appl Genet (2000) 101:814–822

Z. Deng · S. Huang · P. Ling · C. Chen · C. Yu
C.A. Weber · G.A. Moore · F.G. Gmitter Jr.

Cloning and characterization of NBS-LRR class resistance-gene candidate sequences in citrus

Received: 25 October 1999 / Accepted: 27 March 2000

Abstract Numerous disease resistance gene-like DNA sequences were cloned from an intergeneric hybrid of *Poncirus* and *Citrus*, using a PCR approach with degenerate primers designed from conserved NBS (nucleotide-binding site) motifs found in a number of plant resistance genes. Most of the cloned genomic sequences could be translated into polypeptides without stop codons, and the sequences contained the characteristic motifs found in the NBS-LRR class of plant disease resistance genes. Pairwise comparisons of these polypeptide sequences indicated that they shared various degrees of amino-acid identity and could be grouped into ten classes (RGC1–RGC10). When the sequences of each class were compared with known resistance-gene sequences, the percentage of amino-acid identity ranged from 18.6% to 48%. To facilitate genetic mapping of these sequences and to assess their potential linkage relationship with disease resistance genes in *Poncirus*, we developed CAPS markers by designing specific primers based on the cloned DNA sequences and subsequently identifying restriction enzymes that revealed genetic polymorphisms. Three of the amplified DNA fragment markers (designated as 18P33a, Pt9a, and Pt8a) were associated with the citrus tristeza virus resistance gene (*Ctv*), and one fragment (Pt8a) was associated with the major gene responsible for the citrus nematode resistance (*Tyr1*); both genes are from *Poncirus* and of importance to citrus survival and production. These polymorphic fragments were located on two local genetic linkage maps of the chromosome region from *Ctv* to *Tyr1*. These results indicate that resistance-gene candidate sequences amplified with the NBS-derived degenerate primers are valuable sources for developing markers in disease resistance-gene tagging, mapping, and cloning.

Key words Disease resistance genes · Citrus tristeza virus resistance · Citrus nematode resistance · Molecular markers · Genetic mapping

Communicated by M.A. Saghai-Maroof

Z. Deng · S. Huang · P. Ling · C. Chen · C. Yu
F.G. Gmitter Jr. (✉)
University of Florida, Citrus Research and Education Center,
700 Experiment Station Road, Lake Alfred, FL 33850, USA
e-mail: fgg@lal.ufl.edu
Tel.: +1-863-9561151 ext. 301, Fax: +1-863-9564631

C.A. Weber · G.A. Moore
University of Florida, Department of Horticultural Sciences,
Gainesville, FL 32611, USA

Introduction

Cultivated citrus species are susceptible to many diverse pathogens including viruses, viroids, fungi, and bacteria. This susceptibility causes huge losses to the citrus industry, one of the most important fruit crop industries in the world. Breeding for resistance to important diseases has been one of the top priorities in citrus cultivar improvement programs, but improvement has been severely hindered by the lack of effective and efficient selection procedures for disease resistance, and the lack of understanding of the inheritance of resistance traits.

Molecular-marker technologies have developed very rapidly in the last decade, and they have allowed citrus genetists and breeders to locate and map the dominant gene responsible for the citrus tristeza virus resistance (designated *Ctv*) (Gmitter et al. 1996; Deng et al. 1997; Fang et al. 1998) and the major gene for citrus nematode (CN, *Tylenchulus semipenetrans*) resistance in *Poncirus* (designated as *Tyr1*) (Ling et al. 2000). Using similar technologies, markers are being sought to tag other resistance genes in *Citrus* and its relatives. The availability of these markers has greatly facilitated genetic analyses and the utilization of tagged resistance genes in resistant-cultivar development, and it is even allowing the molecular cloning of resistance genes for transfer into desirable cultivars by genetic transformation. So far, the identification of DNA markers for resistance genes has mainly depended on the utilization of the random amplified polymorphic DNA (RAPD) technique.

A number of disease resistance (R) genes have been cloned from several model plant species. Many of these cloned resistance genes appear to encode components of signal transduction pathways, and their protein products share some common structural domains (Baker et al. 1997; Hammond-Kosack and Jones 1997). One of the common domains is the nucleotide-binding site (NBS). Motifs of this domain are well conserved in several R genes, including *Arabidopsis RPS2* (Bent et al. 1994; Mindrinos et al. 1994), tobacco *N* (Whitham et al. 1994), and flax *L6* (Lawrence et al. 1995). Degenerate primers designed from the conserved amino acids in this domain have allowed successful PCR-amplification of multiple DNA sequences from a number of plant species that share striking similarity to the NBS-LRR (leucine-rich repeat) class of resistance genes (Kanazin et al. 1996; Leister et al. 1996; Yu et al. 1996; Seah et al. 1998; Shen et al. 1998, Speulman et al. 1998). These sequences have been called resistance-gene analogs (RGAs) (Kanazin et al. 1996) or resistance-gene candidate (RGC) sequences (Shen et al. 1998). Genetic analyses have associated a number of these sequences to known gene loci that confer resistance to viruses, bacteria, fungi, or nematodes (Kanazin et al. 1996; Leister et al. 1996; Yu et al. 1996; Aarts et al. 1998; Seah et al. 1998; Shen et al. 1998, Speulman et al. 1998). Examples include the very close linkage of two RGA clones from one gene family with the nematode resistance locus *Gro1* of potato (Leister et al. 1996). More intriguingly, some of these sequences appear to be part of the resistance genes themselves. Using one RGA clone as a probe, Yu et al. (1996) identified a soybean genomic clone that contains almost every feature found in tobacco *N* and *Arabidopsis RPS2* genes. Meyers et al. (1998) found that one sequence (*RGC2B*) isolated with RGA probes contains both NBS and LRR domains, and it may be their target *Dm3* resistance gene. These studies seem to indicate that PCR approaches using degenerate primers based on the conserved NBS domains of cloned R genes can provide an attractive strategy to amplify multiple resistance-gene candidate sequences and that these sequences can be developed into molecular markers for use in marker-assisted selection (MAS) or even lead to molecular cloning of new disease resistance genes.

In this study, we cloned and characterized numerous NBS-LRR class resistance-gene candidate sequences from an intergeneric hybrid of *Poncirus trifoliata* and *Citrus grandis*. Specific primers were designed for these sequences and evaluated by PCR-amplification for future use in developing markers for resistance-gene tagging and mapping. RGC-based CAPS (cleaved amplified polymorphic sequence) markers were identified that are closely linked to *Ctv* and *Tyr1*, two important resistance genes in *P. trifoliata*.

Materials and methods

DNA cloning and sequence analysis

Six degenerate primers were used in four combinations to amplify RGC sequences by PCR (Table 1). Primer F11 was one of the two oligonucleotides designed in the sense direction corresponding to the amino-acid sequence GVGKTT found in the P-loop of *N*, *L6* and *RPS2*; primers R11, R16 and R18, were three of the eight oligonucleotides based on the sequence GLPLAL in the anti-sense direction (which is part of a proposed weak hydrophobic region in *N*, *L6* and *RPS2*). Their sequences were: F11, 5´- GG(A/G/T) GT(A/G/T) GGN AA(A/G) AC(A/T) AC; R11, 5´-AGI GC(A/C/T) AGN GGN AGN CC; R16, 5´-AGN GC(A/C/T) AGN GG(C/T) AAN CC; and R18, 5´-AAN GC(A/C/T) AGN GG(C/T) AAN CC. Primers LM637 and LM638 were synthesized from the sequences described by Kanazin et al. (1996). The PCR template used was the genomic DNA of USDA 17–47, an intergeneric hybirid of *C. grandis* and *P. trifoliata,* which was isolated from tender leaves as described by Durham et al. (1992). Amplifications were performed on a MJ PTC-100 thermal cycler (MJ Research) in a 25-μl reaction volume; each reaction contained 50 m*M* Tris-HCl pH 8.3, 2 m*M* $MgCl_2$, 800 μ*M* dNTPs, 25 μ*M* forward and reverse degenerate primers, 150 ng of genomic DNA, and 1 unit of *Taq* polymerase. The initial denaturation was 93°C for 2 min, followed by 42 cycles of 1 min at 92°C, 1 min at 50°C, and 2 min at 72°C. PCR products were separated on agarose or polyacrylamide gels; desired bands were excised, re-amplified, and purified before cloning. Three different vectors, pCR-Script (Stragagene), pGem T and pGem T Easy (Promega), were used in the course of cloning. Clones were characterized by restriction analyses with two or three enzymes (*Hae*III, *Hin*fI or *Taq*I) and classi-

Table 1 Degenerate primers and vectors used in the cloning of citrus RGC sequences. * Number of clones sequenced; clones in parentheses were identical or very similar to the preceding ones, which are indicated with = or ≈, respectively. The percentages in parentheses show the percent amino-acid identities with the preceding clones; Pt20 and Pt23 are identical to Pt18 at the amino-acid level, but differ at five nucleotide bases

Primer combinations	Degenerate primers		Vectors	# of clones*	RGC clones	
	Forward	Reverse			Name	Class
I	F11	R11	pGem T Easy	7	11P31	RGC3
II	F11	R16	pGem T Easy	4	16R1–13, 16R1–19	RGC7
III	F11	R18	pCR-Script	7	18P33 (=18P32), P203	RGC6
					18P34, 18P35	RGC9
IV	LM637	LM638	pGem T	21	Pt6 (=Pt12)	RGC1
					Pt14	RGC2
					Pt3 (≈Pt4, 99%), Pt7 (=Pt21),	RGC4
					Pt18 (=Pt20, Pt23)	RGC5
					Pt8, Pt9 (≈Pt11,99%)	RGC8
					Pt19	RGC10

fied based on the restriction fragment patterns. Representative clones of each class were chosen for sequencing. Double-stranded plasmid DNA was sequenced by the University of Florida's DNA Sequencing Core Laboratory (Gainesville, Fla.) with an ABI 373 automated sequencer and a fluorescently labeled di-deoxy terminator. Sequence editing and analyses were conducted with PC Gene (Intelligent Genetics), DNASIS (Hitachi Software, South San Francisco, Calif.), and Genetics Computer Group (Release 9.1; GCG, Madison, Wis.) software programs. Database searches were performed using Gapped BLAST (Altschul et al. 1997). Phylogenetic analysis was performed using the CLUSTALX package (Thompson et al. 1997). Ten iterations of sequence alignment and tree construction were performed according to Meyers et al. (1999). In each iteration, a neighbor-joining (NJ) tree (Saitou and Nei 1987) was generated and used as the guide tree for the next cycle of alignment. The CLUSTALX default options were used in the initial alignment. Bootstrap analysis was performed in CLUSTALX to evaluate the reliability of the nodes of the phylogenetic trees.

Marker development: primer designing, PCR amplification, and restriction digestion

To develop PCR-based markers for genetic mapping of RGC sequences, sequence regions that were divergent among clones were identified by multiple sequence alignment and subsequently used as bases for designing specific primers. Computer software Oligo 4.0 was used to facilitate the identification of sequence segments with desirable internal stability curves as priming sites and to avoid potential 3′ dimer or hairpin formation (Rychlik 1995). The length of primer oligonucleotides ranged from 21 to 28 bases; this was intentionally varied to obtain similar melting temperatures. Oligonucleotides were synthesized by Operon Technologies. PCR reactions were prepared and run essentially as described previously (Deng et al. 1997). Initially, each pair of primers was evaluated for amplification and polymorphism identification with genomic DNA from 'Thong Dee' pummelo (*C. grandis* L.) and from the intergeneric hybrid USDA 17–40 ['Thong Dee'×Pomeroy trifoliate orange (*P. trifoliata*)]. 'Thong Dee' and '17–40' were the seed and pollen parents, respectively, of the backcross population referred to as the R family that was used in genome mapping (Cai et al. 1994) and CTV resistance-gene mapping (Gmitter et al. 1996; Deng et al. 1997). When restriction digestions were required to reveal polymorphisms between the two parents, 5 µl (for polyacrymide gels) or 10 µl (for agarose gels) of each PCR reaction was incubated 3–16 h with 3 or 5 units of restriction enzymes in a 10-µl or 15-µl vol. PCR products or their digests were separated on 1.4% agarose gels or 6% polyacrylamide gels and detected by staining with ethidium bromide. Markers defined by a pair of specific primers and restriction enzyme digestion are referred as CAPS markers (Konieczny and Ausubel 1993).

Identifying RGC markers linked to CTV and CN resistance genes

BSA (bulked segregant analysis, Michelmore et al. 1991) was employed to assess the potential association of RGC-derived markers with the CTV and CN resistance genes. The resistant and susceptible DNA bulks for CTV resistance were prepared by pooling genomic DNA from eight CTV-resistant or -susceptible individuals of the R family (Gmitter et al. 1996). The CN resistance is a quantitative trait; therefore, genomic DNA from the eight most resistant or susceptible individuals of the CN population (Ling et al. 2000) were combined to construct the resistant and the susceptible bulks. Previously, this pooling procedure allowed the identification of RAPD markers linked to *Tyr1*, the major gene for CN resistance (Ling et al. 2000). PCR-amplification of bulked DNA and restriction digestion of PCR products was performed as described above. The segregation of putative *Ctv*-linked markers was then analyzed among 70 individuals of the R family. Similarly, the segregation of *Tyr1*-associated markers was determined using 63 individuals of the CN population. Based on the segregation data, *Ctv*- or *Tyr1*-linked RGC markers were located onto previously constructed genetic linkage maps (Gmitter et al. 1996; Ling et al. 2000). Marker orders and genetic distances were calculated with MAPMAKER 3.0b (Lander et al. 1987; Lincoln et al. 1992) using a LOD value threshold of 3.0 and the Kosambi mapping function; the validity of the marker orders was tested with the 'RIPPLE' function of MAPMAKER.

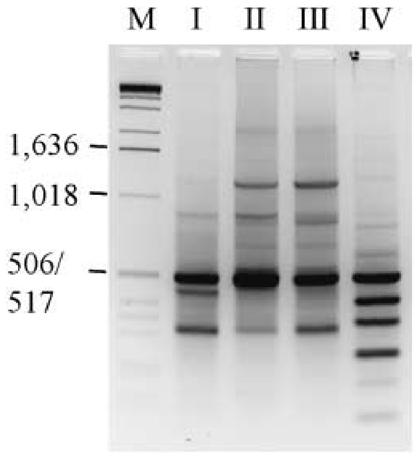


Fig. 1 PCR amplification of genomic DNA of USDA 17–47 using four primer combinations (*I, II, III, IV*). *M*: 1-kb ladder DNA size markers; the size of marker DNA fragments is indicated to the left of the picture in base pairs

Results

PCR amplification and molecular cloning of RGC sequences

The PCR products amplified from the genomic DNA of 17–47 with four degenerate primer combinations are shown in Fig. 1. Primer combinations I, II, and III each generated one major band of around 500 bp in size and a few faint bands. Primer combination IV produced four major bands ranging from 200 bp to 500 bp in size. The approximately 500-bp band from each primer combination was close to the fragment size expected based on the sequences of the *N*, *L6* and *RPS2* genes; therefore, these bands were cloned. Restriction analyses of inserts of randomly selected clones indicated that the 500-bp DNA band from all four primer combinations contained heterogeneous fragments. Clones showing different restriction patterns of insert fragments (occassionlly clones of identical or similar patterns as well) were identified and characterized further by sequencing and sequence analysis.

Sequence analysis

A total of 39 clones were sequenced (Table 1). Searches of the GenBank database using the BLASTN algorithm revealed that two clones were highly similar to chloroplast rDNA, two clones had no hits in the database, five clones had no or only very weak similarity to resistance

```
11P31 -GVGKTTLAREVYNDRSVQDF---KFD--LKAWVCLSD---NFN-VLSISRAILESITS-------APCDLKALNEVQVELKKAVDGK-KILLVVDDVWN
I2C-2 GGLGKTTLAKAVYNDESVKN----HFD--LKAWFCVSE---AYN-AFRITKGLLQEIGS-------IDLVDDNLNQLQVKLKERLKEK-KFLIVLDDVWN
  Pt7 GGVGKTTLAKKLYGDKDVRR----HF---CCAWVSVTQ---DYK-LKDLLLRIIKSFKFKTAL---EDLET--EDDLGRYLHKSLQKH-KYLMVLDDIWD
  Pt3 GGVGKTTLARKLYHNNDVKN----KFD--YCAWVSVSQ---DYK-IKDLLLRIIKSFNIMTAL---EDLETKTEEDLARSLRKSLEAY-SYLMVIDDIWH
 Pt18 GGVGKTTLARKLYHHNDVKH----KFD--CCAWVSVSQ---EYR-TEDLLMRIINSFDIDYP----SNLEKMREEDLERCLYQSLQGY-SYLVVIDDVWQ
 RPM1 GGSGKTTLSANIFKSQSVRR----HFE--SYAWVTISK---SYV-IEDVFRTMIKEFYKEADTQIPAELYSLGYRELVEKLVEYLQSK-RYIVVLDDVWT
16R1-19 -GVGKTTLLKQVNNKFCS-EE--HDFD--VVIWSVVSR---EPN-LMQIQEDIGKRIGFSTD-----SWQRKSLEERASDITNSLKHK-KFVLLLDDIWE
16R1-13 -GVGKTTLLNQVNNKFCGDEQ--HHFD--VVIRSVVSR---EPN-MKQIQEDIGKRIGFSKN-----SWQDKSFEERASDITNTLKHK-KFVLLLDDIWE
  Pt9 GGVGKTTLLTKINNKLLGAP---NGFD--VVIWVVVSK---DLQ-LEKIQEKIGRRIGFLDE-----SWKNGSLEDKASDILRILSKK-KFLLLLDDIWE
  Pt8 GGVGKTTLLTQINNKFLDSRK--DDFD--VVIWVVVSK---DLK-IERIQDDIWKKIGLCDN-----SWRSKSLEDKAVDIFRVLSKK-KFVLLLDDMWK
18P33 -GVGKTTLLKQVNNNFRYQQ---HMFD--VVIWAAVS----------TLQDDIGKRIGFSEDR----NGKEKSLQDKAVDIASILSGK-KFVLLLDDIWE
18P34 -GVGKTTLLRNLNHKFSN-AE--HNFD--RVILVESRT---DVINVETVQFVLKNRPAIPNE-----VWDNKNQQGRAVEIFQRLSQR-RFALLLDDLRG
 RPS2 GGVGKTTLMQSINNELITKG---HQYD--VLIWVQMSR---EFG-ECTIQQAVGARLGLSWD-------EKETGENRALKIYRALRQK-RFLLLLDDVWE
 Pt19 GGVGKTTLVKEIQKQAKEMK----MFD--DVAMAVVSQ---TPT-ITKIQDEIAGWLGVKKL-------PDTDESARASFLWERIKEKQRVLVILDDLWG
    N GGVGKTTIARAIFDTLLGRMDSSYQFD-GACFLKDIKE---NKRGMHSLQNALLSELLR-EK------ANYNNEEDGKHQMASRLRSK-KVLIVLDDIDN
  Pt6 GGVGKTTLARVVYDLIS------HEFE-GSSFLADVREKFENKGSVISFQRQLLFEILKFEK------DSIWNVGDGINILGSRLQHK-KVLLVIDDVVD
 Pt14 GGVGKTTLARFVFDNIS------YQFDDGSSFLANVREVS-QTRGLVALQEQLVSEILLDKN------VKIWDVHKGCHMIRIKLRHK-RVLLVIDDVDE
   L6 GGIGKTTTAKAVYNKIS------SCFD-CCCFIDNIRETQ-EKDGVVVLQKKLVSEILRIDSG----SVGFNNDSGGRKTIKERVSRF-KILVVLDDVDE

11P31 EDYSSWEDLKAPFLVAAPNSKIILTTRHSHVASTMGP--IEHYNLKRLSDEDCWSVFMKHAFEGRDVDG-HQISELYRKKIDGKCGGLPLA 170
I2C-2 DNYNEWDELRNVFVQGDIGSKIIVTTRKDSVALMMG---NEQISMGNLSTEASWSLFQRHAFENMDPMG-HSELEEVGRQIAAKCKGLPLA 169
  Pt7 K--EAWLSLKSAFPEKMNGSRVIITTRNKGVAERLDGQ-TYVHELRFLTPEESWQLFCKKAFHDSIAN---KGLEKLGREMVEKCRGLPLA 168
  Pt3 K--EDWVSLKSAFPENKIGSRVIITTRIKDVAERSDDR-NYVHELRFLRQDESWQLFCERAFRNSKAE---KGLENLGREMVQKCDGLPLA 171
 Pt18 K--ETWESLKRAFPDSKNGSRVILTTRIKEVAERSDER-THVYELPFLRPDNSWKLFCEKAFQSLNAD---EGLEKLGREMLEKCGGLPLA 170
 RPM1 T--GLWREISIALPDGIYGSRVMMTTRDMNVASFPYGIGSTKHEIELLKEDEAWVLFSNKAFPASLEQCRTQNLEPIARKLVERCQGLPLA 178
16R1-19 S-EIDLTKLGVPLQTLDSGSRIVFTTRFEGTCGKMGAH-KNRYKVFCLGDDDAWKLFEGVVGSYALNK--HPDIPKLAEHVARQCHGLPLA 171
16R1-13 F-EIDLTKLGVPLQTLDSGSRIVFTTRFEGTCGKMGAH-KNRYKVFCLRDDDAWKLFEGVVGRYVLNK--HPDIPKFAEDVARQCHGLPLA 172
  Pt9 --RVDLTKVGVPFPNLENKSKIVFTTRFLEICGAIKAH-EF-LKVECLGPEDAWRLFRENLRRDVLDN--HPDIPELARSVAKGCAGLPLA 170
  Pt8 --RVDLTQLGVPLPSPTTASKVVFTTRFVEVCGAMKAH-EY-FKVECLAHEKAWILFQEHVERQTLES--HPDIPELAETVTKECGGLPLA 171
18P33 --RIDLTELGVPLQNLNDGSKIVLTTRSAGVCDQMDSK-K--LEVYSLAHDKAWELFQEMVDRSTLDS--HTSIPELAETLARECGGLPL- 162
18P34 --PINLAEAGVPVQ---NGSKIVYTTIMEDACNVMGDQ-MK-LKVDCLLPDDAWNLFRLMVKDDVLNF--HHDILELAETVADLCGGLPLA 167
 RPS2 --EIDLEKTGVPRPDRENKCKVMFTTRSIALCNNMGAE-YK-LRVEFLEKKHAWELFCSKVWRKDLLE--SSSIRRLAEIIVSKCGGLPLA 168
 Pt19 --RIKLSEVGIPYGKDHRGCNILLTSRSRVVCNQMNAN-K-IVEVGTLTNEESWSRFREVAGPEVDN----LQINPTAREVADGCGGFPLA 166
    N K-DHYLEYLAGDLDWFGNGSRIIITTRDKHLIEKN----DIIYEVTALPDHESIQLFKQHAFGKEVPN---ENFEKLSLEVVNYAKGLPLA 171
  Pt6 I-KQ-LEYLAGKREWFGSGSRIIVTSRDEHLLKTHG--MDEIYKPNELNYHDALQLFNMKAFKIQKPL---EECVQLSEGVLRYVGGLPLA 170
 Pt14 F-DQ-LQALAGQRDWFGLGSRIIITTRDRHLIVRCD--VEDTYMVEKLNYNEALHLFSWKAFRKGHPT---DGYFELSHSMVNYADGLPLA 170
```

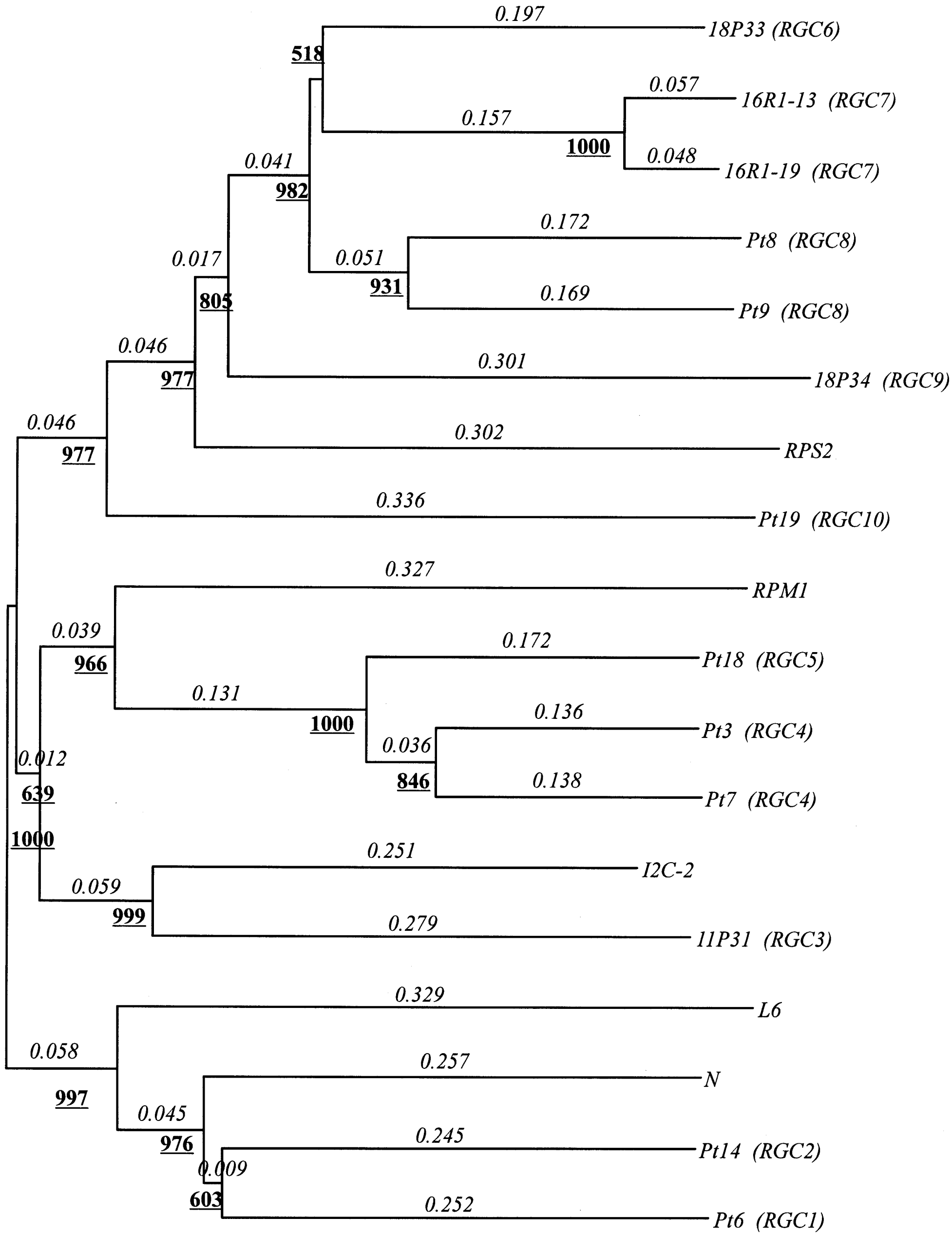
Fig. 2 Alignment of deduced amino acid sequences of 13 citrus RGCs and the NBS domains of five R genes: tobacco *N* (Witham et al. 1994), tomato *I2C-2* (Simons et al. 1998), *Arabidopsis RPM1* (Grant et al. 1995), *RPS2* (Bent et al. 1994; Mindrinos et al. 1994) and flax *L6* (Lawrence et al. 1995). The computer program CLUSTALX was used in the alignment analysis. *Arrows* indicate the priming sites; kinase-2 and kinase-3a motifs are *underlined*

genes; therefore these clones were excluded from further analysis. Eight sequences showed similarity to cloned disease resistance genes at the DNA level, but contained one or more stop codons or frame shifts. The other 22 sequences could be translated to polypeptides without any stop codons, and they showed strong overall similarities to several plant R-gene sequences and many RGA or RGC sequences recently cloned from other plant species using similar PCR-based approaches (BLAST search data not shown). Multiple alignments performed with these

Table 2 Percent amino-acid identities of ten classes of citrus RGCs when compared to each other and to five R genes. Values were calculated using the GCG ‘GAP’ program (at UF/ICBR) with a gap creation penalty =12 and a gap extension penalty =4. The first and last six amino acids of each RGC were excluded in the comparisons. * No sensible values were obtained using the above parameters of GCG ‘GAP’ program

	RGC1	RGC2	RGC3	RGC4	RGC5	RGC6	RGC7	RGC8	RGC9	RGC10	I2C-2	L6	N	RPM1
RGC2	47.1													
RGC3	27.6	22.9												
RGC4	20.9	22.3	30.5											
RGC5	24.5	22.9	34	62.8										
RGC6	–*	21.5	23.8	21.9	24.7									
RGC7	24.1	24.4	20.8	22.1	23.6	57.2								
RGC8	–*	–*	23.8	24.0	23.0	55.3	53.7							
RGC9	–*	–*	22.4	17.6	19.5	44.6	40	38.0						
RGC10	–*	–*	23	20.7	21.6	32.0	31.3	30.2	27.3					
I2C-2	26.1	27.2	44	33.8	30.8	27.8	26.2	27.6	23.2	28.3				
L6	31.9	33.1	–*	21.5	21.8	–*	–*	–*	–*	–*	29.1			
N	45.4	48	–*	29.0	28.8	20.7	22.1	–*	–*	–*	25.2	36.8		
RPM1	–*	–*	25.8	34.8	34.2	20.4	18.6	21.9	–*	23.4	29.0	–*	–*	
RPS2	–*	23.2	22.0	22.4	24.7	40.3	38.9	40.2	34.9	24.2	23.4	–*	–*	25.2

Fig. 3 Phylogenetic tree based on alignment of the deduced amino-acid sequences of 13 citrus RGCs and the NBS domains of five R genes. The tree was constructed using the neighbor-joining method provided in CLUSTRALX. Branch lengths (proportion of amino-acid differences distinguishing classes) are indicated above the lines in *italics*; bootstrap values based on 1000 replications are ***underlined and boldfaced***



citrus RGCs and the five most similar R-gene peptide sequences showed that the similarity was especially high at the three NBS motifs (P-loop, kinase-2, and kinase-3a) (Fig. 2). While the P-loop sequences might have been derived from the degenerate primers F11 or LM638 used in the PCR amplification, the other two motifs should have been amplified from the *Poncirus* or *Citrus* genome. This overall similarity and the existence of the two internal motifs seem to indicate that the 22 sequences belong to the NBS-LRR resistance-gene superfamily.

Pairwise comparisons of the translatable sequences, using the GCG ‘GAP’ program, indicated the percentage indentities at the amino-acid level. Several sequences were highly similar (>99% identity) (Table 1). Using a 70% identity threshold value, these sequences were grouped into ten classes, designated RGC1–RGC10 for simplicity. The percent identities within each of the groups ranged from 73% to 99%, and the identities between groups were generally less than 57% (Table 2). The exception to this was groups RGC4 and RGC5, which shared nearly 63% identity. The percent amino-acid sequence identities of the ten RGC classes compared with previously cloned plant disease resistance genes are listed also in Table 2. RGC groups 1 and 2 were most identical to the *N* gene of tobacco (average identity was 46.7%). RGC3 was most identical (44%) to tomato *I2C-2*, while RGC4 and 5 shared some level of identity with sequences of all five of the previously cloned disease resistance genes; identities ranged from 21.5 to 34.8%. RGC groups 6, 7 and 8 shared their greatest identity (ranging from 38.9 to 40.3%) with *RPS2* from *Arabidopsis*. RGC9 and 10 had a lower similarity with known resistance genes than the other citrus RGC groups.

Phylogenetic analysis was also performed to evaluate further the relationship among citrus RGCs and plant R genes. The deduced amino-acid sequences of 13 representative clones from ten RGC groups and the NBS domains of above five R genes were aligned and a neighbor joining tree was generated from the alignment. The process was performed for ten iterations. This reiterative process was considered necessary as the alignments and trees are mutually dependent (Meyers et al. 1999). A similar tree topology was observed from the ten iterations. Figure 3 shows one typical NJ tree. Pt6 (group RGC1) and Pt14 (group RGC2) together with N and L6 formed one major cluster; the other 11 RGCs from groups RGC3 to RGC10, and RPS2, RPM1 and I2C-2, formed another major cluster. A majority of the tree nodes were supported with >80% of the 1000 replicates in bootstrap analysis.

Marker development and association with disease-resistance genes

To facilitate genetic mapping of the cloned RGC sequences and an assessment of their potential linkages with disease resistance genes in *Citrus* and *Poncirus*, we designed 13 pairs of primers based on the divergent DNA sequence regions (Table 3). Primers were also synthesized for one DNA sequence that showed similarity to R genes at the DNA level, but contained stop codons (clone 11P33). When genomic DNA of 17–40 and 'Thong Dee' was amplified with these 14 pairs of primers, three types of polymorphisms were observed (Table 3). Two primer pairs detected presence/absence polymorphisms. Pt3UP and Pt3LW amplified one band from 17–40, but not from 'Thong Dee'. Pt7UP and Pt7LW amplified one band from 'Thong Dee' but not from 17–40. Primer pair 11P33F and 11P33R amplified one to three bands and revealed a fragment size difference between 17–40 and 'Thong Dee'. The other primers amplified one band from the two test templates that appeared to be of identical or similar sizes, so a panel of restriction enzymes (mainly enzymes with four-base recognition sites) was used to digest the PCR products to reveal polymorphisms. One to three restriction enzymes were able to show polymorphisms for each of the nine primer pairs, thus allowing CAPS markers to be developed for nine RGC sequences. When examined in the R family, these nine markers segregated as co-dominant markers. No polymorphisms were found in the amplification products from primer pair 16R1–13F and 16R1–13R, and primer pair 18P34F and 18P34R.

To demonstrate the potential of associating RGC markers with disease resistance genes, we screened the 14 pairs of primers using two DNA bulks based on the CTV resistance phenotype and another two bulks for CN

Table 3 Polymorphisms identified with RGC-derived primers. [a] Based on the template RGC sequences. [b] Polymorphisms detected between USDA 17–40 and 'Thong Dee' pummelo, the parents of the R family (Gmitter et al. 1996). P/A: presence or absence of DNA bands; numbers in parentheses were the sizes (bp) of polymorphic DNA fragments. Clone 11P33 was derived from degenerate primers F11 and R11 and contained stop codons in its DNA sequence

RGC clones	Primers		PCR product sizes (bp) [a]	Polymorphisms [b]
	Forward	Reverse		
Pt6	Pt6UP: GACTTGATCTCTCATGAATTTGAA	Pt6LW: CACCGACATACCGTAGAACACC	457	*Alu*I (230, 70)
Pt14	Pt14UP: GACAATATCTCTTATCAGTTTGATG	Pt14LW: GAGCTCAAAATAACCATCTGTAG	429	*Hin*fI (430, 280, 150)
11P31	11P31F: TGCTCGGGAGGTCTACAATGACAG	11P31R: CCTCCGCACTTTCCATCAATCTT	477	*Alu*I (75)
Pt3	Pt3UP: CCACAACAATGATGTCAAGAATAA	Pt3LW: GTCCCTTTTCAGCCTTAGAGTTAC	422	P/A
Pt7	Pt7UP: TTACGGCGACAAAGATGTCAG	Pt7LW: TCCCAACTTCTCCAATCCTTTATTA	430	P/A
Pt18	Pt18F: TAAGCATAAATTCGATTGTTGTG	Pt18R: TCCAACTATTATCTGGCCTTAGAA	356	*Taq*I (80), *Alu*I (430, 220)
Pt8	Pt8F: ATTCGCGGAAAGATGATTTTGA	Pt8R: ACACTCTTTCGTCACGGTTTCAG	443	*Taq*I (440)
Pt9	Pt9F: AGCTTCTTGGTGCACCAAATGGTT	Pt9R: CCCTTTAGCTACACTTCTGGCTAGTTCA	446	*Hin*fI (55), *Bfa*I (170), *Nla*III (315)
18P33	18P33F: AAGTCAACAACAACTTCCGCTATCA	18P33R: GGTTTCGGCTAGCTCTGGAATACT	428	*Alu*I (430)
16R1–13	16R1–13F: TGTTAAACCAAGTCAATAACAAATTCAG	16R1–13R: CCATGACACTGACGTGCCACATC	484	None
16R1–19	16R1–19F: TCTGCAGTGAGGAGCATGATTTTGAT	16R1–19R: ATGACACTGACGTGCCACATGCT	455	*Bfa*I (480)
18P34	18P34F: GATCGGGTGATTTTGGTTGAGTC	18P34R: GTCGGCCACAGTTTCAGCAAGTT	411	None
Pt19	Pt19F: AGGAAATTCAGAAGCAAGCAAAAG	Pt19R: ATCCGTCAGCCACCTCTCTT	450	*Hin*fI (170), *Bfa*I (500), *Nla*III (515)
11P33	11P33F: GCAAGCTGCAGGTTGTGGTGTTTA	11P33R: AGGCCGACCTGGTTGAGTTTG	471	Size differences

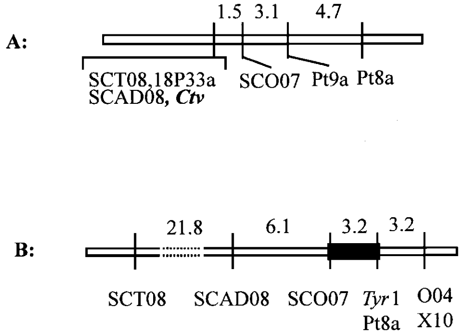


Fig. 4A, B Local linkage maps of the *Ctv* and the *Tyr1* region. Maps **A** and **B** were constructed using segregation data from the R family and the CN family, respectively. Marker orders were generated using the functions 'GROUP' and 'COMPARE' in MAPMAKER EXP 3.0b and tested using the 'RIPPLE' function with its default value (window size 3 and log-likelihood threshold value 2). The distance between markers is shown above the linkage groups in Kosambi centiMorgans. SCT08, SCAD08 and SCO07 are SCAR markers developed previously (Deng et al. 1997); 18P33a, Pt8a and Pt9a are CAPS markers derived from RGCs; O04 and X10 are RAPD markers amplified with Operon decamer primer O04 and X10 (Ling et al. 2000), respectively. Three markers, SCT08, SCAD08 and 18P33a were co-localized with *Ctv* based on their segregation data from the R family. Citrus nematode resistance is a quantitative trait; the major gene for the resistance, *Tyr1,* was mapped to the region defined by markers SCO07 and Pt8a (Ling et al. 2000), as indicated with a *solid box* in map **B**

resistance. Prior to this study, CTV resistance and CN resistance were characterized in the R family and the CN family, respectively, and localized genetic maps were constructed. Bulked segregant analyses revealed that three pairs of primers (Pt8F and Pt8R, Pt9F and Pt9R, and 18P33F and 18P33R) produced three polymorphic DNA fragments (Pt8a, Pt9a, and 18P33a) between CTV-resistant and -susceptible bulks. Further analyses of 70 R family individuals confirmed that these polymorphisms were linked to *Ctv*; a genetic map was constructed using their segregation data and the data from three previously developed SCAR markers (SCT08, SCAD08, and SCO07). 18P33a co-segregated with markers SCT08 and SCAD08 and with *Ctv* on this map (Fig. 4A). Fragment Pt8a was also polymorphic between the two DNA bulks for CN resistance. It was added onto the genetic map developed by Ling et al. (2000) (Fig. 4B), in between marker SCO07 and co-segregating RAPD markers O04 and X10, which identified the *Tyr1* region. SCAR markers SCO07, SCAD08 and SCT08 were placed on the map to allow for comparisons between maps.

Discussion

RGC sequences are associated with CTV and CN resistance genes

The use of PCR approaches with degenerate oligonucleotide primers designed from the NBS region of cloned disease resistance genes has led to the cloning of resistance gene-like sequences in several plant species (Kanazin et al. 1996; Leister et al. 1996; Yu et al. 1996; Aarts et al. 1998; Seah et al. 1998; Speulman et al. 1998). Correlation or co-segregation of some of these sequences with known disease resistance gene loci has been documented. Improvement of disease resistance has been one of the top priorities in citrus breeding; development of DNA markers can significantly expedite this process. We have been vigorously searching for markers associated with resistance to several important citrus pathogens including CTV and CN. For this purpose, the value of the above-mentioned PCR approach was evaluated in this study. Out of 39 clones that resulted from PCR with four primer combinations, 22 RGC sequences were identified. Three polymorphic DNA fragments amplified with RGC sequence-derived primers were linked to *Ctv* in *Poncirus*. Fragment Pt8a was also polymorphic between Swingle citrumelo and LB 6–2, the donor and recurrent parents, respectively, of the CN backcross family. This marker was mapped to the middle of a previously identified region that contains *Tyr1* for CN resistance. These results indicate that RGC sequences are valuable sources for marker development in disease resistance-gene tagging and mapping. The availability of several specific PCR-based markers (SCAR and CAPS markers) for *Ctv* and *Tyr1* should facilitate the selection and introgression of these resistance genes into new varieties.

The cloned citrus RGC sequences showed 18–48% amino-acid identities to cloned R genes. The percent identities between these RGCs varied significantly as well. Using a 70% threshold value, the sequences were grouped into ten classes. The percent identity within each group was generally greater than 73% and the identity between groups was less than 57%. Shen et al. (1998) used percent amino-acid identity in grouping lettuce RGCs. In lettuce, members of the same gene families shared at least 50% identity and different families had less than 40% identity. Compared to their criteria, a much higher threshold value was used in this study. Phylogenetic analysis results seem to support this grouping. Preliminary genetic mapping data seemed also to indicate that members of the same groups, such as Pt8 and Pt9, shared similar locations on chromosomes (unpublished data). Other approaches used to classify RGC sequences include cross-hybridization, but the results vary with the stringency of post-hybridization washes (Shen et al. 1998). This approach was not used to classify citrus RGCs.

Meyers et al. (1999) performed extensive analyses of the NBS domains of the plant NBS-LRR class R genes and a vast number of RGCs and found that they could be classified into either TIR (Toll/Interleukin-1 receptor homology) or non-TIR groups. In their analysis, N and L6 belong to the TIR group, while RPS2, RPM1 and I2C-2, fall into the non-TIR group. Our data indicate that Pt6 (RGC1) and Pt14 (RGC2) formed a major cluster with N and L6, and they have an aspartic acid residue (D) at the final residue position of the kinase-2 motif that is often seen in the TIR group. Most of citrus RGCs were clustered with RPS2, RPM1, and I2C-2 and have a tryptophan residue (W), indicating that they might belong to the non-TIR group. Attempts are being made to obtain the upstream sequence information; when this becomes available, it should assist in the classification of these citrus RGCs regarding TIR or non-TIR grouping.

The *Ctv*-*Tyr1* region contains a major cluster of resistance genes

Previously we reported the clustering of *Ctv* and *Tyr1* based on one common SCAR marker (SCO07; Deng et al. 1997). The current data from RGC marker Pt8a and SCAR marker SCAD08 support this observation. In addition, three RGC sequences were mapped to this region. The 18P33a fragment was cloned and found to be over 95% identical to the original RGC; it contained an uninterrupted open reading frame similar to R genes. Recently, Fang et al. (1998) found that two of their *Ctv*-linked RAPD marker fragments (OPJ07$_{650}$ and OPC19$_{960}$) were very similar to *RPS2* (Bent et al. 1994; Mindrinos et al. 1994) and *Cf2* (Dixon et al. 1996), respectively. We cloned the two RAPD fragments from USDA 17–47 (progeny of Pomeroy trifoliate orange) and observed the same similarity with the two R genes. Considered together, these data seem to indicate that a major cluster of resistance-gene candidate sequences may exist in the *Ctv*-*Tyr1* chromosomal region, and that it may consist of at least five NBS-LRR class RGC sequences, one LRR class sequence, and two functional resistance genes of unknown nature. Such clustering of R genes has been found in other plants (Kanazin et al. 1996; Meyers et al. 1998; Shen et al. 1998) and seems to be a common distribution feature in plant genomes. Disease resistance genes in citrus were not genetically located or mapped until recently (Gmitter et al. 1996), and the clustering of R genes in citrus was poorly understood. Therefore, further genetic and physical mapping of the *Ctv*-*Tyr1* genomic region and cloning of some of these resistance genes would provide important information on the evolution and function of disease-resistance genes in *Poncirus* and *Citrus*. Toward this end, numerous bacterial artificial chromosome (BAC) clones containing RGC sequences have been isolated. Sequences downstream from and upstream of some of the cloned RGC regions have been examined and found to contain the other motifs that R genes possess, such as LRRs (unpublished data).

Structural re-arrangement(s) in the *Ctv*-*Tyr1* chromosomal region?

Based on the strong similarity of *P. trifoliata* strains shown by molecular fingerprinting (Fang et al. 1997), we assumed that most strains might share similar gene (or marker) order and genetic distances in the *Ctv*-*Tyr1* region. An attempt was made to generate an integrative map of this region by using segregation data from the R family and the CN family, the former being derived from a large-flowered strain ‘Pomeroy’, and the latter being the progeny of Swingle citrumelo from an unrecorded *P. trifoliata* strain. Markers SCAD08, SCT08 and 18P33a all co-segregated in the R family (and were mapped within 4 cM in a large population, unpublished data), but fell far apart on the map generated from the CN population. This significant disagreement in genetic distances was unexpected and seemed to indicate that chromosomal re-arrangement(s) might have occurred in or around the *Ctv*-*Tyr1* region. Other lines of available data seem to support this explanation. Segregation of CTV resistance and associated markers in the R family was significantly skewed toward the recurrent parent type (‘Thong Dee’; Gmitter et al. 1996). Fang et al. (1998) analyzed the segregation of CTV resistance in several families of different origins and nature, and their data showed that the segregation of CTV resistance was significantly skewed in the cross involving ‘Pomeroy’, as well. We are analyzing more markers located in this region in several diverse families to compare genetic maps. If the re-arrangement could be confirmed, it raises interesting questions about the nature and size range of the structural change and whether or not the change involves *Ctv*. Is this gene identical in different *P*. *trifoliata* strains (large flowered vs small flowered)? Molecular cloning of this resistance gene is being vigorously pursued; when successful, it should shed light on these questions.

Acknowledgements The authors thank Ms. Marjorie Wendell for her assistance with the research. This project was supported in part by grants from the USDA NRICGP (#9600748), USDA/National Citrus Research Council (#98012205), the Florida Citrus Production Research Advisory Council (#942–27), and a USDA-ARS/UF-IFAS Specific Cooperative Agreement. All experiments described in this paper comply with the current laws of the United States of America. Florida Agricultural Experiment Station Journal Series No. R-06785.

References

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

Altschul SF, Thomas L, Madden A, Schäffer A, 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

Baker B, Zambryski P, Staskawicz B, Dinesh-Kumar SP (1997) Signaling in plant-microbe interactions. Science 276:726–733

Bent AF, Kundel 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

Cai Q, Guy CL, Moore GA (1994) Extension of the linkage map in *Citrus* using random amplified polymorphic DNA (RAPD) markers and RFLP mapping of cold-acclimation-responsive loci. Theor Appl Genet 89:606–614

Deng Z, Huang S, Xiao S, Gmitter FG Jr (1997) Development and characterization of SCAR markers linked to the citrus tristeza virus resistance gene from *Poncirus trifoliata*. Genome 40:697–704

Dixon MS, Jones DA, Keddie JS, Thomas CM, Harrison K, Jones JDG (1996) The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell 84:451–459

Durham RE, Liou PC, Gmitter FG Jr, Moore GA (1992) Linkage of restriction fragment length polymorphisms and isozymes in *Citrus*. Theor Appl Genet 84:39–48

Fang DQ, Roose ML, Krueger RR, Federici CT (1997) Fingerprinting trifoliate orange germplasm accessions with isozymes, RFLPs, and inter-simple sequence repeat markers. Theor Appl Genet 95:211–219

Fang DQ, Federici CT, Roose ML (1998) A high-resolution linkage map of the citrus tristeza virus resistance gene region in *Poncirus trifoliata* (L.) Raf. Genetics 150:883–890

Gmitter FG, Jr., Xiao SY, Huang S, Hu XL, Garnsey SM, Deng Z (1996) A localized linkage map of the citrus tristeza virus resistance gene region. Theor Appl Genet 92:688:695

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 Plant Mol Biol 48:575–607

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

Konieczny A, Ausubel FM (1993) A procedure for mapping *Arabidopsis* mutations using co-dominant ecotype-specific markers. Plant J 4:403–410

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

Ling P, Duncan LW, Deng Z, Dunn D, Hu X, Huang S, Gmitter FG Jr (2000) Inheritance of citrus nematode resistance and its linkage with molecular markers. Theor Appl Genet 100: 1010–1017

Lincoln S, Daly MJ, Lander E (1992) Constructing genetic maps with MAPMAKER EXP3.0, 3rd edn. Whitehead Institute Technical Report, Cambridge, Massachusetts

Meyers BC, Chin DB, Shen KA, Sivaramakrishnan S, Lavelle DO, Zhang Z, Michelemore 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, Sivaramakrishnan S, Sobral BW, Young ND (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucelotide-binding superfamily. Plant J 20: 317–332

Michelemore 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 genomic regions by using segregating populations. Proc Natl Acad Sci USA 88:9828–9832

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

Rychlik W (1995) Selection of primers for polymerase chain reaction. Mol Biotechnol 3:129–134

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

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

Shen KA, Meyers BC, Islam-Faridi M. N, 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

Simons G, Groenendijk J, Wijbrandi J, Reijans M, Groenen J, Diergaarde P, Lee TV, Bleeker M, Onstenk J, Both M, Haring M, Mes J, Cornelissen B, Zabeau M, Vos P (1998) Dissection of the *Fusarium I2* gene cluster in tomato reveals six homologs and one active gene copy. Plant Cell 10:1055–1068

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

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24:4876–4882

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

Yu YG, Buss GR, Saghai-Maroof MA (1996) Isolation of a superfamily of candidate disease-resistance genes in soybean based on a conserved nucleotide-binding site. Proc Natl Acad Sci USA 93:11751–11756