Mol Gen Genomics (2005) 273: 240–251
DOI 10.1007/s00438-004-1104-7

ORIGINAL PAPER

Hans van Leeuwen · Jordi Garcia-Mas · María Coca
Pere Puigdoménech · Amparo Monfort

Analysis of the melon genome in regions encompassing TIR-NBS-LRR resistance genes

Received: 30 July 2004 / Accepted: 9 December 2004 / Published online: 12 April 2005

Abstract Plant genomes contain numerous genes (R-genes) that play a role in initiating defence measures against their particular pathogens. Defence mechanisms controlled by R-genes have been the focus of extensive research over the past several years. The majority of the R-genes described so far belong to a super-family of genes (150–600 members) that encode proteins with a nucleotide binding site (NBS), some leucine-rich repeats (LRR) and an N-terminal domain that shows similarity to the Toll and Interleukin-1 receptors (TIR) or a N-terminal coiled-coil (CC) domain. Analysis of four regions of the melon (*Cucumis melo*) genome, including two sequenced BACs, identified 14 TIR–NBS–LRR genes. Known disease resistance genes have been mapped in three of these regions. Transcriptional expression was detected for predicted genes that are possibly involved in defence responses to pathogen attack. TIR–NBS–LRR genes appear to be clustered in the melon genome. They contain all the conserved motifs that have previously been described for their counterparts in other species, although differences were also detected. The results presented here may contribute to a better understanding of the genomic distribution and evolution of this group of resistance gene homologues and their variability.

Communicated by M.-A. Grandbastien

H. van Leeuwen · J. Garcia-Mas · M. Coca
P. Puigdoménech · A. Monfort
Laboratori de Genética Molecular Vegetal,
Consorci CSIC-IRTA,
Barcelona, Spain

H. van Leeuwen · M. Coca · P. Puigdoménech
Departament de Genética Molecular,
Jordi Girona 18, 08034 Barcelona, Spain

J. Garcia-Mas · A. Monfort (✉)
Departament de Genética Vegetal,
Institut de Recerca i Tecnologia Agroalimentáries (IRTA),
Ctra. de Cabrils s/n, 08348
Cabrils, Barcelona, Spain
E-mail: amparo.monfort@irta.es
Tel.: +34-937-507511
Fax: +34-937-533954

Introduction

Plants use a variety of mechanisms to defend themselves against pathogen attack. In many cases, plant disease resistance genes (R-genes) have been shown to confer resistance against a pathogen in accordance with the "gene-for-gene" model originally described for the flax–flax rust interaction by Flor (1956). Many different types of R-genes, encoding proteins with different functional domains, have been characterised in a variety of species (reviewed in Hammond-Kosack and Parker 2003). The largest class of functionally defined R-genes encode products that have a nucleotide binding site (NBS) domain and a leucine-rich repeat (LRR) domain. Resistance gene homologues (RGHs) of the NBS–LRR class occur in large numbers in plant genomes. For example, in the *Arabidopsis* genome nearly 150 NBS–LRR RGHs have been identified (Meyers et al. 2003). This class of R-genes is divided into two groups depending on whether the N-terminal domain is a Toll/Interleukin-1 receptor (TIR) domain (Whitham et al. 1994) or a coiled-coil (CC) domain (Bent et al. 1994; Mindrinos et al. 1994). The specificity of the R-gene for different races of the same pathogen can be determined by the TIR domain (Luck et al. 2000) or the LRR domain (van der Hoorn et al. 2001).

RGHs of the NBS–LRR type are often organised in clusters in plant genomes, as has been demonstrated for *Arabidopsis thaliana* (thale cress) (Meyers et al. 2003), *Oryza sativa* (rice) (Zhou et al. 2004), *Lycopersicon esculentum* (tomato) (Parniske et al. 1997), *Glycine max* (soybean) (Ashfield et al. 2003), *Medicago truncatula* (barrel medic) (Zhu et al. 2002) and *Cucumis melo* (melon) (van Leeuwen et al. 2003), amongst others. Recombination events that generate new R-genes, and

thus potentially provide protection against different pathogens, may involve mainly recombination between NBS–LRR genes from the same cluster (Meyers et al. 2003; Zhu et al. 2002), although recombination between diverse NBS–LRR genes has also been proposed (Richly et al. 2002).

The Cucurbitaceae family contains species that are important crops world wide and are second only to Solanaceae in their economic value among horticultural crops. The cultivated melon is susceptible to a variety of diseases, so the identification and mapping of resistance genes can contribute to identifying melon varieties of increased agronomic value. The melon genome is compact, being comparable in size to that of rice, and only three times larger than that of *Arabidopsis*. Microcolinearity has been demonstrated between a region of the melon and *Arabidopsis* genomes (van Leeuwen et al. 2003), thus allowing the use of comparative genomics approaches. Conserved synteny would make this species a valuable tool for increasing our understanding of plant genomes and of the structure, distribution and evolution of plant disease resistance genes. Four RGH sequences were previously isolated from melon (Garcia-Mas et al. 2001). These MRGHs were placed on the linkage map of melon, and BACs located in these regions were identified. In this study, we present the analysis and comparison of four regions of the melon genome that together contain 14 TIR–NBS–LRR RGHs. Expression studies were performed on predicted genes in these regions in order to assess the possible involvement of these genes in responses to pathogen infection.

Materials and methods

Screening of the PIT92 BAC library for MRGHs

A PCR screening for MRGH18 and MRGH21 was performed using DNA pools of the PIT92 BAC library (van Leeuwen et al. 2003). We replicated the library (sixty 384-well plates) using a 96-pin replication tool into sixty 96-well plates with four clones in each well, containing 200 μL of medium, as described by Zhang et al. (1996). After overnight growth at 37°C, a 50 μL aliquot was taken from each well of one plate, added to 2.5 mL of LB containing chloramphenicol 12.5 µg/mL and grown overnight at 37°C. Glycerol stocks of each plate culture (60) were stored at −80°C. The remainder of the bacterial culture was used for the isolation of BAC DNA as described by Zhang et al. (1996), resulting in 60 DNA pools of the PIT92 library, each containing 384 BAC clones. DNA from five pools each was mixed to obtain 12 DNA superpools. The remaining bacterial culture in the 96-well plates was used to create row DNA pools and column DNA pools for each plate. Each row DNA pool contained 48 BAC clones, while each column DNA pool contained 32 BAC clones.

PCRs with MRGH18- and MRGH21-specific primers were performed with the DNA superpools as template, followed by a second round of PCR with the DNA pools from the positive DNA superpools. PCR primer sequences were: MRGH18-U (5′-GTTAGAGAAGCTTCAGAGCA-3′), MRGH18-L (5′-CAACAAGTTCTGAAAGGTCC-3′), MRGH21-U (5′-GATGGATGTTATTTTCTGGAC-3′) and MRGH21-L (5′-CTACAACTTGGCTACAGAGAT-3′). PCRs were carried out in a total volume of 25 μL, and contained 10 mM Tris–HCl (pH 8.3), 50 mM KCl, 2.5 mM $MgCl_2$, 100 µM dNTPs, each primer at 0.25 µM, 1.75 U of Taq DNA-polymerase (Roche Expand High Fidelity) and 0.5 μL of BAC miniprep DNA. PCR cycling conditions were: 2 min at 94°C, 35 cycles (20 s at 94°C, 20s at 54°C, 30 s at 72°C), 5 min at 72°C. Amplified sequences of the expected size were purified with the GFX PCR DNA and Gel Band Purification Kit (Amersham Pharmacia Biotech), cloned into the pGEM-T vector and sequenced. After identifying a positive DNA pool we performed a third round of PCR with the row DNA pools and column DNA pools. We then performed a fourth round of PCR with the four BAC clones from the original PIT92 BAC library to identify the positive BAC.

Construction of physical maps

The MRGH18, MRGH21 and MRGH7 contigs and physical maps were constructed as previously described for the MRGH63 contig (van Leeuwen et al. 2003). BACs belonging to the MRGH7 contig were detected during the screening of the library for MRGH21 and by comparison of sequenced amplified fragments.

Genetic mapping of the MRGHs

For MRGH18 we developed two markers from the U and SP6 BAC-end sequences of the clones 24L19 and 52B5, respectively. The primer pair for the 24L19D SSR (single sequence repeat) marker (24L19D-F: 5′-ATTGGACCAAAATCAATTTGC-3′; 24L19D-R: 5′-AAAGTGTCATTTGTACTAGAG-3′) amplified a fragment of 275 bp (from PI 161375) or 257 bp (from Piel de Sapo). The reaction and cycling conditions were as described for MRGH18 and MRGH21, except that the annealing temperature was 52°C. The primer pair for the 52B5SP6 CAPS (cleaved amplified polymorphic sequence) marker (52B5SP6-F: 5′-TACTCTCTACTAAATGGGAC-3′; 52B5SP6-R: 5′-GCAATAGTGTATTGCGTTAC-3′) amplified a fragment of 472 bp from both PI 161375 and Piel de Sapo. The reaction and cycling conditions were as described for 24L19D SSR. The PCR product was digested with the *Xmi*I (Fermentas), according to the manufacturer's instructions. *Xmi*I only cleaves the fragment amplified from PI 161375, resulting in fragments of 346 and 126 bp. Digestion products were separated on a 2% w/v agarose gel.

For MRGH7, we developed a CAPS marker based on the SP6 BAC-end of the 31L3 clone. With the primers 31L3SP6-F (5′-GGATCCAATCCTAAGATAAGG-3′) and 31L3SP6-R (5′-TTCTTAAGACCCACTTGCAAC-3′) we amplified a fragment of 272 bp from both PI 161375 and Piel de Sapo. The reaction and cycling conditions were as described for 24L19D SSR, except that the annealing temperature of 56°C. The PCR product was digested with *Nco*I (Fermentas), according to the manufacturer's instructions. *Nco*I only cleaves the fragment amplified from PI 161375, resulting in fragments of 231 and 41 bp. Digestion products were separated on a 2% w/v agarose gel.

The double haploid lines (DHLs) from the melon genetic mapping population (Monforte et al. 2004) were genotyped with these markers, allowing us to place the new markers on the melon genetic map.

BAC subcloning, sequencing and assembly

A subclone library of the 31O16 BAC, consisting of 1,440 clones with an average insert size from 1,200 bp to 1,400 bp, and plasmid DNA from these clones were provided by GATC Biotech. For sequencing with the Terminator Cycle Sequencing Ready Reaction kit (Perkin Elmer) using the T3 and M13 forward primers, 8 μL of plasmid DNA was used. The sequencing reactions were processed on an ABI PRISM 377. Assembly of the subclones into one contig was done with the pregap and gap4 software in the STADEN package (Staden 1996), and the Phred-Phrap-Consed software (Gordon 1998). Remaining gaps were bridged by PCR with flanking primers, and sequencing of the amplified regions.

Sequence analysis

The sequence of the BAC 31O16 was analysed to find genes and repetitive elements as described previously (van Leeuwen et al. 2003). Localisation of *Arabidopsis* sequences homologous to the melon sequences used the TBLASTN program (Altschul et al. 1997) at the TAIR BLAST web site (http://www.arabidopsis.org). Additional analysis for localisation of repeated sequences included the use of the JDotter software (Brodie et al. 2004; http://athena.bioc.uvic.ca/pbr/jdotter/). For the phylogenetic analysis of the MRGHs, multiple protein sequence alignment was performed with the ClustalW 1.8 program (Thompson et al. 1994) at the BCM SearchLauncer web site (Smith et al. 1996; http://searchlauncher.bcm.tmc.edu/). A distance matrix was calculated using the PHYLIP (Felsenstein 1989) PROTDIST program with the PAM matrix setting. The tree was generated by the neighbour-joining method using the PHYLIP NEIGHBOR program. The tree was visualised with the program TreeView (Page 1996).

Expression studies

The expression of the predicted genes *MRGH5, MRGH63* and *Drzf* from the BAC 60K17 (AF499727), and *MRGH11, MRGH21, PDI2* and *PDI4* from the BAC 31O16 (AY582736) was studied as follows. Total RNA was isolated from young leaves of *C. melo* L. cv. Piel de Sapo 4 days after infection with Tobacco Rattle Virus (TRV), and from plants that had not been infected with the virus, as described by Logemann et al. (1987). Virus inoculations were performed using *Agrobacterium* infiltration. The leaf material was kindly provided by Montserrat Martin of the IRTA (Cabrils, Spain). The isolated RNA was treated with DNase I (Roche Diagnostics) according to the manufacturer's instructions, extracted with phenol–chloroform–isoamylalcohol extraction and finally precipitated with ethanol. For cDNA synthesis, we used the M-MLV reverse transcriptase (Invitrogen) according to the manufacturer's instructions and a primer consisting of dT_{32}, together with 2-μg aliquots of DNase-treated RNA. For each gene five PCRs were performed with different template DNAs: (1) cDNA from non-infected plants, (2) gDNA from *C. melo* L. Piel de Sapo, (3) the product of a reaction carried out in the absence of reverse transcriptase, (4) cDNA from TRV-infected plants 4 days after inoculation, and (5) a control reaction without DNA template. The following primer pairs were designed and used for RT-PCRs: *MRGH5*-F (5′-ACCATCAACCCAGTATGCTCTTC-3′) and *MRGH5*-R (5′-ATAGGTAGGAGATGAACCTGCTG-3′), *MRGH63*-F (5′-TCATCCGTCTTTACCTTCCTCAC-3′) and *MRGH63*-R (5′-CCACCTTTACAACATTCCTGCAC-3′), *Drzf*-F (5′-CAAGTAACAAGGCGTTGAATCGG-3′) and *Drzf*-R (5′-TCCCCTGTCCAAGTTCTCGGC-3′), *MRGH11*-F (5′-GAAACCTCAATAACCCATGTTCC-3′) and *MRGH11*-R (5′-ATAGGTAGGAGATGAACCTGCTG-3′), *MRGH21*-F (5′-TTACTCTCAAGGGTTGAGAG-3′) and *MRGH21*-R (5′-TTACCTCAATGTATCGTTGAGCC-3′), *PDI2*-F (5′-TGATGCCTACCCTTTCACTC-3′) and *PDI2*-R (5′-AGTAAAAGGACCAACCAGATCCC-3′), *PDI4*-F (5′-TGATGCTTATCCTTTTACTGCTG-3′) and *PDI4*-R (5′-TTCTCATCCTTAAAAGGCAATGC-3′).

PCRs were performed in 15-μL volumes containing 2 mM $MgCl_2$, 0.1 mM dNTPs, 0.25 μM primers, 1 U Taq DNA polymerase (Roche Diagnostics) and approximately 1.5 μg cDNA. PCR cycling conditions for *MRGH5, MRGH63, MRGH11, MRGH21, PDI2, PDI4* and *Drzf* were: 2 min at 94°C, 35 cycles (30 s at 94°C, 30 s at 58°C, 40 s at 72°C), 5 min at 72°C.

Results

Physical contigs encompassing RGH clusters

Four families of melon resistance gene homologues (MRGHs) have been previously identified in *C. melo* by

PCR amplification with degenerate primers based on conserved regions of known resistance genes (R-genes) (Garcia-Mas et al. 2001). Screening of the PIT92 library with the MRGHs identified clones that were assembled into contigs which were genetically mapped to four different locations in the melon genome (Fig. 1). The MRGH63 contig, and the sequence analysis of the BAC to which it belongs (60K17), was described previously (van Leeuwen et al. 2003). The BAC 60K17 maps to linkage group 4 in a region that contains three genetically mapped disease resistance genes: *Vat, Pm*-w and *Agr* (Garcia-Mas et al. 2001). Screening of the BAC library for MRGH18 and MRGH21 was performed by PCR with DNA pools from the PIT92 library (van Leeuwen et al. 2003). Using CAPS and SSR markers, developed from the BAC-end sequences, we mapped the MRGH18 contig to linkage group 8 of the melon map (Oliver et al. 2001). MRGH21 was previously mapped to linkage group 7 in a region with two disease resistance genes (*Fom*-1 and *Prv*) and one quantitative trait locus (QTL) for resistance against the Cucumber Mosaic Virus (CMV) (Garcia-Mas et al. 2001). The 31O16 BAC from this contig was sequenced and characterised. Screening of the PIT92 library with MRGH21 also resulted in the discovery of two BACs with a new melon resistance gene homologue (MRGH7) that is nearly identical to MRGH21. We mapped MRGH7 on linkage group 7 near MRGH21, using a CAPS marker developed from the BAC-end 31L3-SP6. In total, together with the MRGHs that we identified previously (Garcia-Mas et al. 2001), we now have fourteen MRGH sequences, of which nine proved to be full-length genes.

Sequencing and characterisation of the BAC 31O16

Gene content prediction

The BAC 31O16 was found to be 159,477 bp long and it was sequenced with a redundancy of x4. The sequence of 31O16 has been deposited in GenBank under the Accession No. AY582736. Sequence annotation (Fig. 2; Table 1) revealed 21 putative genes, including eight MRGHs (described below), and different types of repetitive elements (described below). The majority of the predicted genes are similar to known genes with a wide variety of putative functions, while three predicted genes are hypothetical with no homologues in the databases (Table 1).

We detected four putative protein disulfide isomerase (PDI) genes clustered together, but on both strands, in the 31O16 BAC: *PDI1, PDI2, PDI3* and *PDI4* (Fig. 2). The degree of identity between the proteins encoded by these genes ranged from 36% (between PDI1 and PDI2) to 77% (between PDI2 and PDI3). BLASTP analysis of the PDI3 sequence revealed highest similarity with a PDI from *Quercus suber* (GenBank AAS02080). The PDI3 and AAS02080 proteins show 58% identity and 73% similarity.

Analysis of repetitive sequences

Transposons

In the first intron of *MRGH10*, we discovered a sequence similar to a LINE transposable element. This element (LINE-1, approximately 5.6 kb long; Fig. 2) contains two ORFs encoding a Gag protein (ORF1) and endonuclease and reverse transcriptase domains (ORF2). The 3′ terminus of the element is marked by a poly(A) tail. We discovered two truncated sequences with similarity to *LINE-1* (*LINE-2, LINE-3*; Fig. 2). We also discovered two putative retrotransposable elements of the Ty1/Copia type (Fig. 2). The LTR of *Copia-2* was not found directly upstream of ORF1, but was separated from it by *Copia-1* and *LINE-3*, and was inverted (Fig. 2). For *Copia-1* we detected the target site duplication flanking the element. Two additional repetitive sequences were found flanking the *Copia-1* element (indicated by the

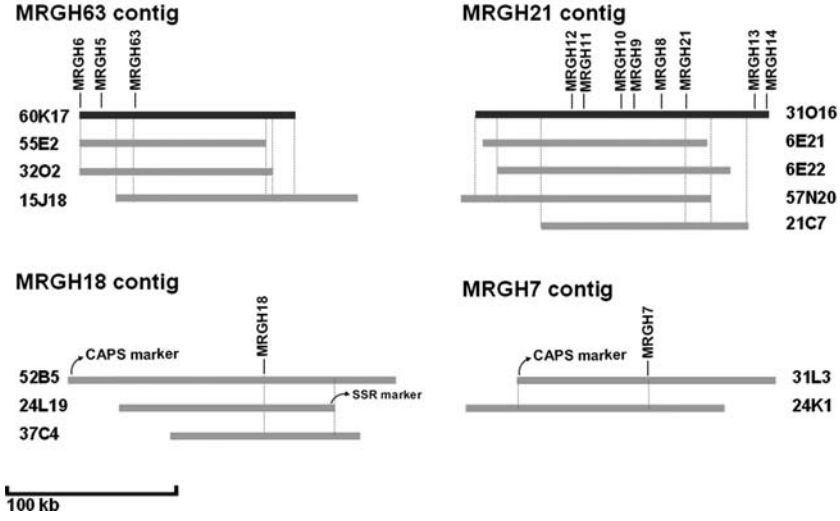

Fig. 1 Physical contigs encompassing the resistance gene clusters. BACs shown in *dark grey* were fully sequenced. The type and location of markers used for mapping of the MRGH18 and MRGH7 contigs are indicated. The MRGH63 contig is derived from linkage group 4, the MRGH21 and MRGH7 contigs map to linkage group 7, and the MRGH18 contig is on linkage group 8

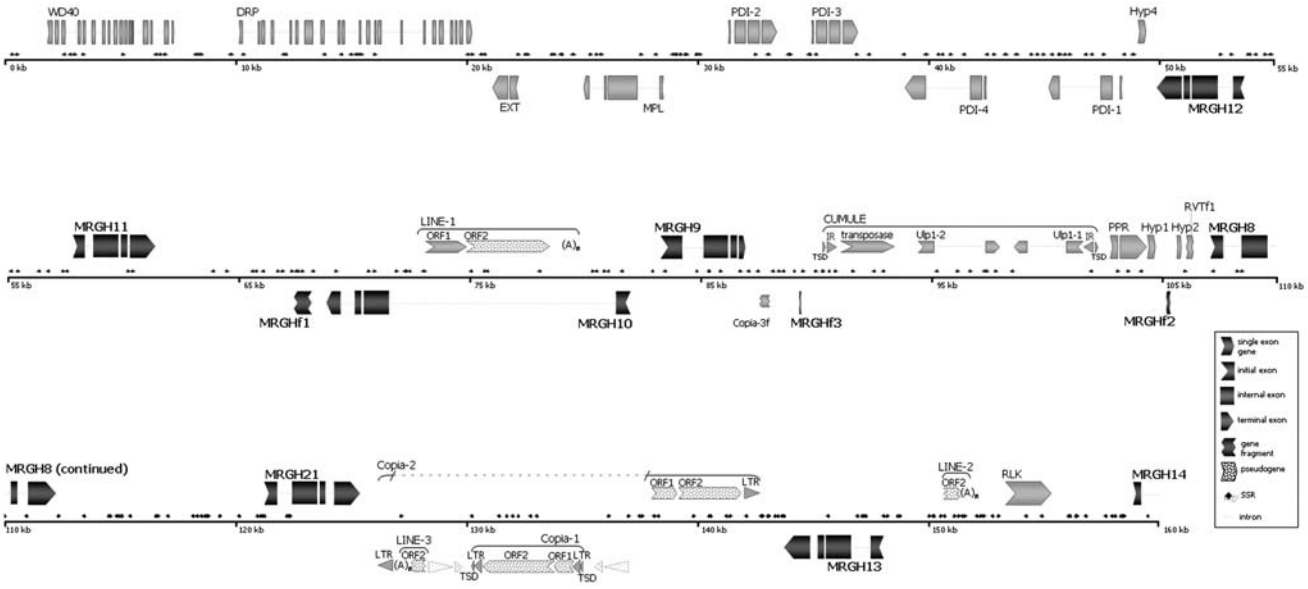

Fig. 2 Annotation of the 31O16 BAC clone. The clone was sequenced and analysed as described in Materials and methods. The sizes of the exons, introns and other elements are shown to scale, except for the simple sequence repeats (SSRs) which are exaggerated for clarity. MRGHs are indicated in *dark grey*. Abbreviations: *ORF* open reading frame; *IR* inverted repeat; *TSD* target site duplication; *LTR* long terminal repeat. Gene names are explained in Table 1

triangles in Fig. 2). These sequences do not show similarity to known repetitive sequences. Another type of transposable element (CUMULE, for *Cucumis Mutator*-like element) was found downstream of *MRGH9*, with similarity to MULE elements in Arabidopsis (Le et al. 2000).

Simple-sequence repeats (SSRs)

The distribution of different types of SSRs in the 31O16 BAC is presented in Table 2. The SSRs in the 31O16 sequence were mostly located in intergenic regions and introns, as in the case the BAC 60K17 (van Leeuwen et al. 2003). In general, for short mono-, di-, tri- and tetranucleotide repeats we see the same types of SSR

Table 1 Predicted genes found in the BAC 31O16 BAC and their putative functions

Gene(s)[a]	Pfam domains	Predicted function
WD40	PF00400 (WD40)	Signal transduction
DRP	PF00350 (Dynamin_N), PF01031 (Dynamin_M), PF00169 (PH), PF02212 (GED)	Energy production
EXT	PF03016 (Exostosin)	Cell wall synthesis
MPL	PB006826 (Pfam-B_6826)	Unknown function; similar to CH1 (*Homo sapiens*)
PDI-1, 2, 3, 4	PF00085 (Thioredoxin), PF03107 (DC1)	Protein folding
Hyp-1, 2	–	Hypothetical proteins
Hyp-4	PF00650 (CRAL_TRIO) (E-value 0.4)	Hypothetical protein, PF00650; belongs to transfer proteins
MRGH-8, 9, 10, 11, 12, 13, 14, 21; MRGHf-1[b], 2[b], 3[b]	PF01582 (TIR); PF00931 (NB-ARC); PF00560 (LRR)	Signal transduction in plant defence against pathogens
PPR	PF01535 (PPR)	Unknown function; widespread in plants.
RLK	PF00139 (Lectin_legB), PF00138 (Lectin_legA), PF00069 (Pkinase)	Signal transduction
Transposable elements		
LINE-1,2[b], 3[b]	PF03372 (Exo_endo_phos), PF00078 (RVT)	Retrotransposable element of the LINE type
Copia-1[b], 2[b], 3[b]	PF00098 (zf-CCHC), PF00665 (rve)	Retrotransposable element of the Copia/Ty1 type
CUMULE	PF03108 (MuDR), PF04434 (SWIM), PF02902 (Peptidase_C48)	Transposable element of the CUMULE type

[a] *DRP* dynamin-related protein, *EXT* exostosin family, *MPL* membrane protein-like, *PDI* protein disulfide isomerase, *Hyp* hypothetical protein, *MRGH* melon resistance gene homologue, *PPR* PPR repeat protein, *RLK* receptor lectin kinase, *LINE* long interspersed nuclear element, *CUMULE Cucumis Mutator*-like element *WD40*, WD40/transducin repeat protein; [b] Pseudogenes or gene fragments

Table 2 Analysis of simple sequence repeats (SSRb) in BAC 31O16

SSR type[a]	No. SSR	kb/SSR
Mononucleotide	**120**	**1.3**
T/A	120	1.3
Dinucleotide	**71**	**2.2**
TA/AT	37	4.3
TC/GA	23	6.9
TG/CA	11	14.5
Trinucleotide	**77**	**2.1**
AAG/CTT	22	7.2
TTA/TAA	21	7.6
AAC/GTT	13	12.3
TGA/TCA	7	22.8
TCC/GGA	5	31.9
AGC/GCT	4	39.9
ACC/GGT	2	79.7
CCG/CGG	2	79.7
TAC/GTA	1	159.5
Tetranucleotide	**9**	**17.7**
TAAA/TTTA	4	39.9
TTTG/CAAA	2	79.7
CTTT/AAAG	1	159.5
TTAA/TTAA	1	159.5
TCCC/GGGA	1	159.5
Pentanucleotide	**3**	**53.2**
GAAAA/TTTTC	2	79.7
TGACA/TGTCA	1	159.5
Hexanucleotide	**2**	**79.7**
AAATCT/AGATTT	1	159.5
TTTTAA/TTAAAA	1	159.5
SSR length (bp)		
8–11	223	0.7
12–19	53	3.0
>20	6	26.6

[a] The minimum length for SSRs (in repeats) was 8(mononucleotides), 4(dinucleotides), 3(trinucleotides), 3(tetranucleotides), 3(pentanucleotides) and 3(hexanucleotides)

occurring in similar numbers in 31O16 (Table 2) and 60K17 (van Leeuwen et al. 2003). The SSR density in 31O16 (one SSR per 0.57 kb) is lower than that in 60K17 (one per 0.35 kb). The majority of the SSRs in 31O16 are of the T/A and TA/AT types, and their frequency is comparable to that found in the *Arabidopsis* genome (Casacuberta et al. 2000). The density of SSRs with repeat lengths of 12–19 bp is one per 3.0 kb, lower than that in the BAC 60K17 (van Leeuwen et al. 2003) and the rice genome (Temnykh et al. 2001).

Sequence analysis of MRGHs

We have identified eight new MRGHs in the 31O16 BAC, all belonging to the TIR–NBS–LRR type. The MRGHs are interspersed with other genes and transposable elements, spanning a region of approximately 100 kb. All MRGHs consist of four exons with introns that are conserved in position and phase. The introns in the MRGHs clearly separate the TIR, NBS and LRR domains (Fig. 3). Intron phases can be classified based on the position of the intron with respect to the reading frame of the gene (Sharp 1981). With one exception, the first intron in all MRGHs is in phase 2; the exception is *MRGH8* where intron 1 is in phase 1. The second and third introns of all MRGHs are in phase 0. The gene architecture is comparable to that of the NBS–LRR genes in *Arabidopsis* (Meyers et al. 2003). The MRGHs contain all the known motifs characteristic for the TIR, NBS and LRR domains (Meyers et al. 1999, 2003).

In the N-terminal TIR domain, several of the MRGHs contain sequences similar to the $MA(S)_n$ consensus that has been found in a variety of highly expressed genes (Sawant et al. 2001), including TIR–NBS–LRR genes from *Arabidopsis* (Meyers et al. 2003). The TIR domains contained four conserved motifs: TIR1–4 (Fig. 3). In the TIR1 and TIR3 motifs we detected insertions in some of the MRGHs that had not been detected in other species.

In the NBS domain we also detected some novel insertions. In the conserved Kin-2 motif, the aspartic acid (D) residue which is conserved in the TIR group of NBS–LRR proteins (Young 2000) is not present in all MRGHs. In MRGH13 and MRGH7 we detected a single amino acid change [from aspartic acid (D) to asparagine (N) that had been found in the predicted products of several TIR–NBS–LRR genes from other species. However, in MRGH9, MRGH10 and MRGH11 multiple mutations were detected in the Kin-2 motif. The conserved GLPL motif, on which degenerate primers are often based, is followed by a 'GSSL' consensus that is also present in other species. Between the NBS domain and the LRR domain we detected a sequence similar to the NL linker (Meyers 2003), which resides in exon 3. At the 3′ end of this exon we also detected two LRR motifs overlapping with the NL linker.

The LRR domain is characterised by leucine-rich repeats found C-terminal to the NBS domain in many R-gene products (Jones and Jones 1997). Sequence alignment of the MRGHs (Fig. 3) shows the modular structure of the LRR domain, with inserted or deleted copies of the LRR unit. We detected a variable region in the LRR domain in several MRGHs that does not seem to encode LRRs. The C-terminal non-LRR sequence contains a novel 'DCVSL' consensus sequence in many MRGHs (MRGH10 is an exception). Most characterized NBS–LRR genes encode proteins directed to the cytoplasm. Exceptions include the *Arabidopsis RPM1* gene (Boyes et al. 1998) and the soybean *KR1* gene (He et al. 2003), which most probably encode peripheral plasma membrane proteins. For the KR1 protein transmembrane segments have been predicted in silico. Using the SMART and Pfam databases we detected potential transmembrane segments in MRGH63, MRGH10 and MRGH8. However, the motif suggested for MRGH8 gave a score that was below the normal detection threshold.

Transcriptional expression studies

Using the RT-PCR technique we studied the transcriptional expression of the predicted genes *MRGH5,*

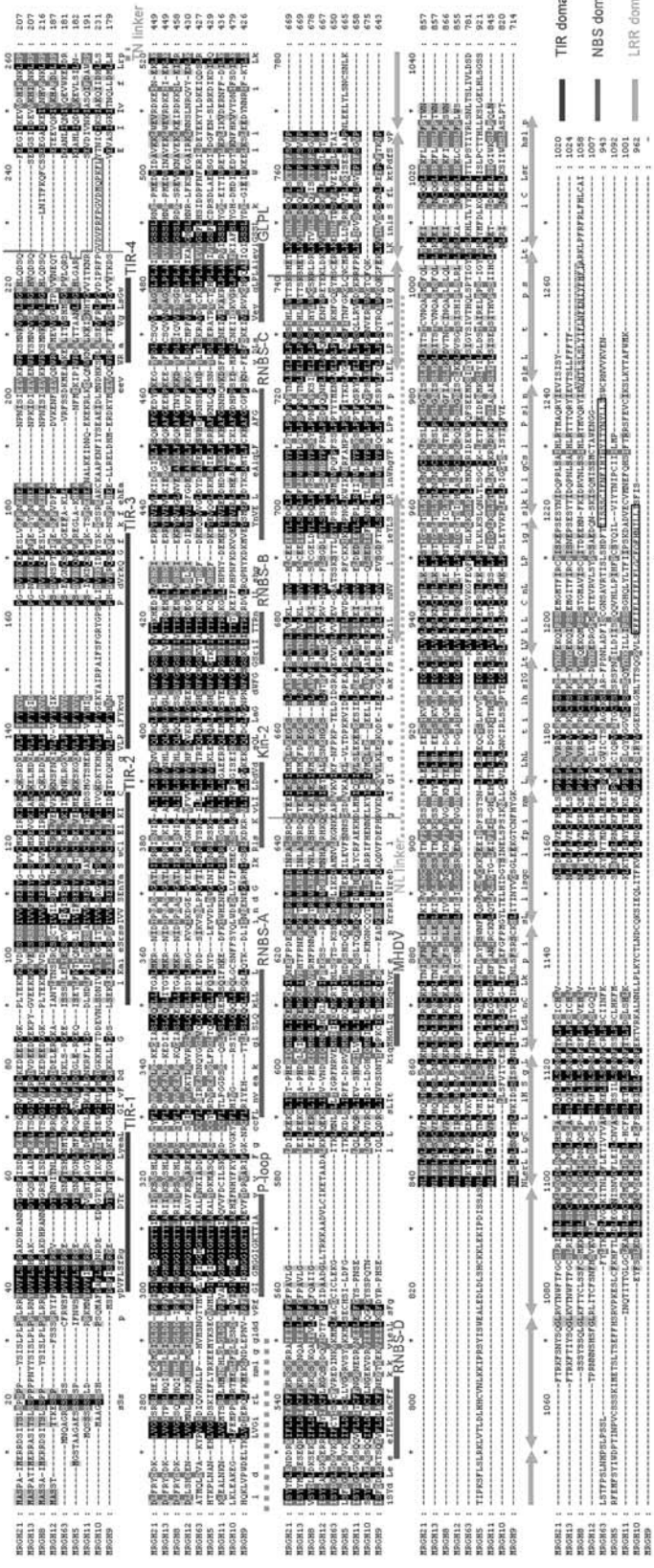

Fig. 3 Sequence alignment of melon MRGHs. TIR–NBS motifs are defined as described in Meyers et al. (1999, 2003), LRRs were detected with Pfam. The *vertical lines* indicate exon boundaries; transmembrane regions detected by TMHMM and SMART are depicted as *rectangles*. The *dotted rectangle* marks a putative transmembrane region for which the prediction score was below the set threshold

MRGH63 and *Drzf* (BAC 60K17), and *MRGH11, MRGH21, PDI2* and *PDI4* (BAC 31O16) in young melon leaves 4 days after infection with TRV. *MRGH11* and *MRGH21* (Fig. 4), and *MRGH5* (results not shown) are expressed in both non-infected and infected plants, but the expression level seems to be stronger in infected plants. Expression of *MRGH63* could not be detected (Fig. 4). Transcriptional expression in both non-infected and infected plants was observed for *Drzf* (Fig. 4), and for *PDI2* and *PDI4* (results not shown). In addition to these experiments, we also confirmed the expression of *MRGH18* in non-infected plants by RT-PCR (results not shown). The expression of the genes *MRGH11, MRGH21* and *MRGH63* in response to the viral pathogen TRV or the fungal pathogen *Fusarium oxysporum* f. sp. *melonis* was also analysed by Northern hybridisation, and multiple transcripts were detected (data not shown). The three genes were expressed in both non-infected and infected plants, although expression levels were higher in the infected plants (data not shown).

Phylogenetic analysis of MRGHs

The isolation of RGH sequences from many plant species, and the availability of the complete sequence of the Arabidopsis genome (The *Arabidopsis* Genome Initiative 2000) with nearly 150 RGHs (Meyers et al. 2003), allows comparative phylogenetic analysis of the MRGHs. For the phylogenetic tree construction we used all MRGHs that have the NBS domain sequence, and selected R-genes and RGHs from other plant species (Fig. 5). From *Arabidopsis* we used NBS–LRR sequences from each of the different branches of the phylogenetic trees constructed for the TIR–NBS–LRR and CC-NBS-LRR sequences by Meyers et al. (2003; http://niblrrs.ucdavis.edu/). All MRGHs belonging to the TIR–NBS–LRR group cluster together in the phylogenetic tree. We can distinguish three subgroups of MRGHs: (1) MRGH9, MRGH11 and MRGH10, (2) MRGH12, MRGH8, MRGH7, MRGH13 and MRGH21, and (3) MRGH63, MRGH5, MRGH4 and MRGH18. Comparing the MRGHs with RGHs from other species we can see that they are more similar to sequences from species that are evolutionarily less remote from Cucurbitae (e.g. soybean) than to Arabidopsis RGHs.

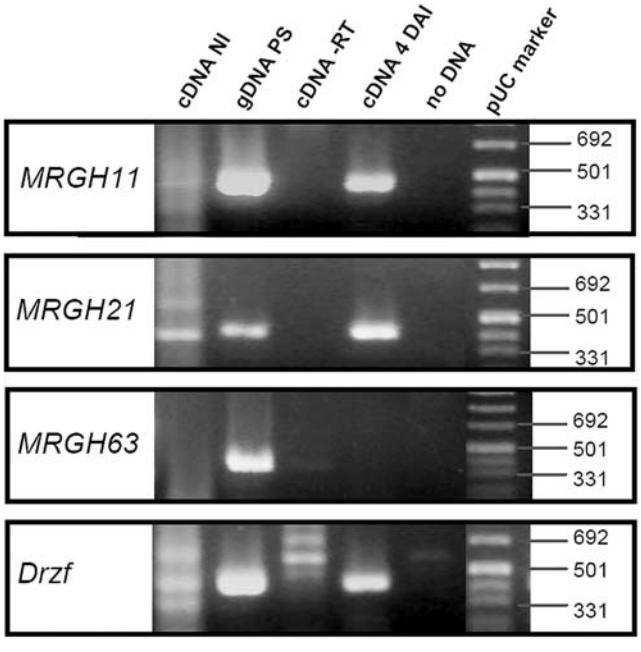

Fig. 4 Transcriptional analysis by RT-PCR of predicted genes that may be involved in defense responses against pathogens. Infection of plants, RNA extraction and PCR conditions are described in Materials and methods. The templates used for RT-PCR were as follows: cDNA NI, RNA isolated from non-infected plants; gDNA PS, genomic DNA from cv. Piel de Sapo; cDNA-RT, PCR was carried out in the absence of reverse transcriptase; cDNA 4 DAI: RNA isolated from plants 4 days after infection with Tobacco Rattle Virus (TRV); no DNA, no template was added

Synteny with *Arabidopsis*

The existence of conserved synteny between melon and *Arabidopsis* has been demonstrated previously (van Leeuwen et al. 2003). The *Arabidopsis* genome was analysed for conserved synteny with the 31O16 BAC sequence, but no single *Arabidopsis* region showed complete synteny to the entire BAC. However, when the 31O16 BAC sequence was first compared with just the TIR–NBS–LRR clusters in *Arabidopsis* (Meyers et al. 2003; http://niblrrs.ucdavis.edu/) a single large cluster on *Arabidopsis* chromosome 5 was found to contain several of the additional non-RGH genes that are present in the 31O16 BAC. This region (At5g44870–At5g46665; 820 kb) contains 24 TIR–NBS–LRR RGHs, one extososin-like gene, one WD40-like gene, three PPR-like genes, one LINE-like and four copia-like elements, and two CUMULE-like transposable elements. Each of these genes is present in multiple copies in the *Arabidopsis* genome (>30 as detected with BLASTP at the TAIR web site at http://www.arabidopsis.org) and this makes the conserved synteny less significant. However, the TIR–NBS–LRR genes in the melon genome are clustered in a similar way to that found in the *Arabidopsis* genome.

Discussion

Physical maps of RGH clusters in the melon genome

Physical maps were constructed encompassing four regions of the MRGHs. The BACs that make up these regions were identified by PCR based screening using DNA pools from the PIT92 BAC library (van Leeuwen et al. 2003). The usefulness of this screening strategy, first demonstrated with human BAC libraries (Asakawa

Fig. 5 Phylogenetic analysis of MRGHs. The sequences of the region between the NBS and GLPL motifs of the MRGHs (indicated by the *grey bars*) and R-genes and RGHs from other species were subjected to phylogenetic analysis (see Materials and methods for details). The sequences used in the comparison and their sources were the following. *Arabidopsis thaliana*: At4g33300 (Q9SZA7), At4g26090-RPS2 (Q42484), At1g15890 (Q9LMP6), At1g61190 (O22727), At1g12280 (P60838), At3g07040-RPM1 (Q39214), At3g46530-RPP13 (Q9M667), At5g43470-RPP8 (Q8W4J9), At1g59218 (NP_176153), At5g45260-RRS1 (Q9FH83), At5g45250-RPS4 (NP_199338), At5g36930 (NP_198509), At5g40100 (NP_198826), At4g16950-RPP5 (CAB80966), At5g46450 (NP_199457), At3g44480 (NP_190034), At2g14080 (NP_179024), At5g41550 (NP_198970), At1g63750 (NP_176562); *Arachis cardenasii*: Ac.RGH (AAN85379); *Brassica napus*: Bn.RPM1.1A (AAC99464); *Cicer arietinum*: Ca.CP3 (AAF36333), Ca.CP6 (AAF36336); *Cucumis melo*: Cm.MRGH18 (CAB88868), Cm.MRGH4 (CAB88871), Cm.MRGH21 (AY582736), Cm.MRGH63 (AAO45749), Cm.MRGH5 (AAO45748), Cm.MRGH8 (AY582736), Cm.MRGH9 (AY582736), Cm.MRGH10 (AY582736), Cm.MRGH11 (AY582736), Cm.MRGH12 (AY582736), Cm.MRGH13 (AY582736), Cm.MRGH7 (AY582142); *Glycine max*: Gm.RLG1 (T08819), Gm.RLG2 (T08820), Gm.RLG8 (T08835); *Gossypium barbadense*: Gb.RGH (AAO89153); *Helianthus annuus*: Ha.NBSR3 (AAD00827); *Lens culinaris*: Lc.RGH (CAD56845); *Lycopersicon esculentum*: Le.I2C (AAB63275), Le.MI (AAC67238); *Lactuca sativa*: Ls.RGC4a (AAC02205); *Linum usitatissimum*: Lu.L6 (AAA91022), Lu.M (AAB47618); *Malus prunifolia*: Mp.RGH (AAM77263); *Manihot esculenta*: Me.RCa2 (AAO38214); *Nicotiana glutinosa*: Ng.N (AAA50763); *Oryza sativa*: Os.Pib (BAA93618); *Phaseolus vulgaris*: Pv.OB5 (AAK50047); *Pinus radiata*: Pr.RGH (T08068); *Pisum sativum*: Ps.RGA2 (AAD52715): *Pyrus communis*: Pc.RGH (CAE46654); *Solanum tuberosum*: St.331 (T07766); *Triticum aestivum*: Ta.CRE3 (AAC05834); *Vitis amurensis*: Va.RGH (AAR08831); *Zea mays*: et al. 1997), was thus confirmed in melon. The use of pools for PCR screening resulted in fewer false positives and was much easier and faster than the radioactive hybridisation strategy previously employed (van Leeuwen et al. 2003).

Three of the resistance gene clusters were found to map near regions where known disease resistance genes have been genetically mapped previously, and this makes the identified MRGHs possible molecular markers for these genes.

Melon genome characteristics compared to *Arabidopsis*

The complete sequence of the 31O16 BAC from the *MRGH21* region allows comparison with the sequence of the BAC 60K17 which includes *MRGH63* (van Leeuwen et al. 2003). The 31016 BAC contains 21 predicted genes, two copia-like retrotransposable elements, three LINE-like retrotransposable elements, one MULE-like transposon, and 282 SSRs. The BAC contains a cluster of eight TIR–NBS–LRR RGHs.

Besides the cluster of TIR–NBS–LRR genes, we found a cluster of four protein disulfide isomerase (*PDI*) genes that might also be involved in disease resistance responses. One plant–pathogen interaction in which a PDI may be involved is the infection of wheat by the fungal pathogen *Mycosphaerella graminicola*. A *PDI* gene is rapidly induced after infection, as demonstrated by Ray et al. (2003), who also reviewed the possible role of PDIs in plant–pathogen interactions. In animals and

plants, PDI proteins are induced under stress conditions and provide enhanced chaperone activity for the assembly of polypeptides. The antioxidant properties of PDI may play a role in limiting the potential for cell damage by reactive oxygen species that are produced during the hypersensitive response in plant–pathogen interactions. Therefore, this genome region might harbour different genes having a function in defence. The four melon *PDI* genes are most similar to a *PDI* gene from *Quercus suber* that is linked to resistance against the bacterial pathogen *Phytophthora cinnamomi* (Coelho 2004).

The 31O16 BAC has a lower gene density than 60K17, which can be due to the greater average gene length and larger number of exons per gene. The 31O16 BAC also has fewer SSRs than 60K17, but contains a larger number of transposable elements. The gene density in both of these regions of the melon genome is lower than the average for *Arabidopsis*, and the genes in melon have larger exons and introns, and thus the average gene is longer. The percentage of transposable elements in the two melon BACs varies significantly and it is therefore difficult to compare this figure with that for the whole genome of *Arabidopsis* at this stage.

Detailed analysis of the MRGHs

In the four regions of the melon genome examined here (Fig. 1) 14 RGHs of the TIR–NBS–LRR class have been identified. Six of these (*MRGH21, MRGH18, MRGH4, MRGH63; MRGH5* and *MRGH6*) we reported previously (Garcia-Mas et al. 2001; van Leeuwen et al. 2003), while eight others (*MRGH7, MRGH8, MRGH9, MRGH10, MRGH11, MRGH12, MRGH13* and *MRGH14*) are presented for the first time in this work. The full length gene and predicted protein sequences are now available from nine MRGHs, and this has allowed a detailed comparison on both DNA and protein levels, of the TIR, NBS and LRR domains. In general, the MRGHs have a conserved modular structure, with the TIR, NBS and LRR domains (Fig. 3), as was also reported for the RGHs present in the Arabidopsis genome (Meyers et al. 2003).

The TIR domains of all MRGHs contain all the conserved motifs described earlier for the Arabidopsis RGHs, but in some of these motifs insertions are found in some of the MRGHs. The insertions of 3–16 amino acids in the conserved TIR domain (Meyers et al. 2003) in some MRGHs (see Result and Fig. 3), which seem to be specific for the melon genes, might alter their pathogen specificity (Ellis et al. 1999; Luck et al. 2000). Among the MRGHs we have detected five genes that do not contain the consensus Kin-2 motif (Young 2000). Analysis of the NBS–LRR RGHs in the Arabidopsis genome revealed that the TIR–NBS–LRR class has two predominant versions of the GLPL motif: NLPL or G NPL [the residues that differ from the consensus (Meyers et al. 2003) are underlined]. In the MRGHs we detected five different versions of the GLPL motif: GLPL, G HPL, RLPL, DLPL or G HP Q. The deviations from the consensus sequence of the Kin-2 and GLPL motifs should be considered when designing degenerate PCR primers for amplification of RGHs. The degenerate primers that we used for cloning MRGHs (Garcia-Mas et al. 2001) do not fully correspond to the TIR–NBS–LRR specific GLPL motifs (Meyers et al. 2003) and this might explain why we obtained relatively few MRGHs in those experiments.

In the C-terminal domain we detected putative transmembrane segments in MRGH63, MRGH10 and MRGH8. The presence of two predicted C-terminal transmembrane segments in the TIR–NBS–LRR protein KR1 from soybean suggests that it is located in the plasma membrane (He et al. 2003). Although the *Arabidopsis* CC–NBS–LRR disease resistance protein RPM1 does not contain a predicted transmembrane segment, it was found to be a peripheral plasma membrane protein, probably residing on the cytoplasmic face of the plasma membrane (Boyes et al. 1998). These results suggest that some, if not all, of the NBS–LRR proteins are localised in the cytoplasm as peripheral plasma membrane proteins.

Expression studies

The results of our expression studies using RT-PCR confirm the validity of our gene predictions as well as the transcriptional expression of the genes *MRGH5, MRGH18*, *MRGH11, MRGH21, Drzf, PDI2* and *PDI4* in TRV-infected plants and in non-infected plants. Expression of *MRGH11, MRGH21* and *MRGH63* was detected by Northern hybridisation in plants infected with TRV or the fungal pathogen *Fusarium oxysporum* f. sp . *melonis*. The results suggest that these genes respond to viral infection, but a response specific for fungal attack could not be detected. Further studies will be required to elucidate the possible roles of these genes in defence against pathogen attack.

MRGHs belong to the TIR-NBS-LRR group

The phylogenetic analysis of the MRGHs and NBS–LRR proteins from other species clusters the MRGHs within the NBS–LRR subclass that contains the TIR domain. In general we can see that the MRGHs from one genomic region are more related to each other than to MRGHs from other genomic regions, indicating that clusters arose as a result of local duplications. The MRGHs from the MRGH21 region and MRGH7, which maps close to MRGH21, fall into two distinct groups. MRGH12 groups with MRGH8, MRGH21 and MRGH13 in the phylogenetic analysis, while in the genome it is separated from these MRGHs by three other MRGHs (see Fig. 2). This suggests the involvement of chromosomal rearrangements, as was already demonstrated to occur in melon (van Leeuwen et al.

2003). For the two BACs that were sequenced, tandem duplications of the TIR–NBS–LRR genes appear to have increased their number in the melon genome, as is the case in *Arabidopsis* (Leister 2004). The phylogenetic tree shows that the MRGHs are more similar to RGHs from species other than *Arabidopsis*. These species, e.g. soybean, are evolutionarily closer to melon than is *Arabidopsis* (Soltis et al. 1999). Our results suggest that this family of genes may more readily form groups by species than by other criteria that could indicate a functional relationship. This may be the consequence of the non-reciprocal recombination events that occur between these gene clusters which are proposed to be the source of their variability.

MRGH clusters similar to those found in *Arabidopsis*

The majority of the NBS–LRR genes are organised in clusters in the genome (reviewed by Michelmore and Meyers 1998). This is most probably also the case for the melon genome, as indicated by the sequence organization of the BACs 31O16 BAC presented in this work and 60K17 (van Leeuwen et al. 2003). With the complete genome sequence of *Arabidopsis* available, the location of nearly 150 NBS–LRR genes has been studied in detail (Meyers et al. 2003). The melon TIR–NBS–LRR genes seem to be clustered in the melon genome in a manner similar to that observed in the *Arabidopsis* genome. Our results suggest tandem duplication as the mechanism responsible for the increase in the number of this class of genes in the melon genome.

Acknowledgements The authors would like to thank Mercé Miquel from the IBMB Sequencing Service for the DNA sequencing, Montserrat Martin for providing the TRV infected melon leaves, and David Lightfoot for providing much useful discussion and helpful comments on the manuscript. Funds for this project were provided by Grant No. 2FD97-0286-(02–01) from the Comisión Interdepartamental de Ciencia y Tecnologia (CICYT) and Grant No. 2000FI-00316 to Hans van Leeuwen from the Comissió Interdepartamental de Recerca i Tecnologia (CIRIT). This work was carried out within the context of the Centre de Referencia de Biotecnologia (CERBA) of the Generalitat de Catalunya.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Asakawa S, Abe I, Kudoh Y, Kishi N, Wang Y, Kubota R, Kudoh J, Kawasaki K, Minoshima S, Shimizu N (1997) Human BAC library: construction and rapid screening. Gene 191:69–79

Ashfield T, Bocian A, Held D, Henk AD, Marek LF, Danesh D, Penuela S, Meksem K, Lightfoot DA, Young ND, Shoemaker RC, Innes RW (2003) Genetic and physical localization of the soybean *Rpg1-b* disease resistance gene reveals a complex locus containing several tightly linked families of NBS–LRR genes. Mol Plant Microbe Interact 16:817–826

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

Boyes DC, Nam J, Dangl JL (1998) The *Arabidopsis thaliana* RPM1 disease resistance gene product is a peripheral plasma membrane protein that is degraded coincident with the hypersensitive response. Proc Natl Acad Sci USA 95:15849–15854

Brodie R, Roper RL, Upton C (2004) JDotter: a Java interface to multiple DotPlots generated by Dotter. Bioinformatics 20:279–281

Casacuberta E, Puigdoménech P, Monfort A (2000) Distribution of microsatellites in relation to coding sequences within the *Arabidopsis thaliana* genome. Plant Sci 157:97–104

Coelho AC (2004) Identification of molecular markers in *Quercus suber* linked to resistance to *Phytophthora cinnamomi*. PhD Thesis, FERN, Universidade do Algarve

Ellis JG, Lawrence GJ, Luck JE, Dodds PN (1999) Identification of regions in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity. Plant Cell 11:495–506

Felsenstein J (1989) PHYLIP—Phylogeny Inference Package (Version 3-2). Cladistics 5:164–166

Flor HH (1956) The complementary genic systems in flax and flax rust. Adv Genet 8:29–54

Garcia-Mas J, van Leeuwen H, Monfort A, de Vicente MC, Puigdoménech P, Arús P (2001) Cloning and mapping of resistance gene homologues in melon. Plant Sci 161:165–172

Gordon D, Abajian C, Green P (1998) Consed: a graphical tool for sequence finishing. Genome Res 8:195–202

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

He CY, Tian AG, Zhang JS, Zhang ZY, Gai JY, Chen SY (2003) Isolation and characterization of a full-length resistance gene homolog from soybean. Theor Appl Genet 106:786–793

Jones DA, Jones JDG (1997) The role of leucine-rich repeat proteins in plant defences. Adv Bot Res 24:90–167

Le QH, Wright S, Yu Z, Bureau T (2000) Transposon diversity in *Arabidopsis thaliana*. Proc Natl Acad Sci USA 97:7376–7381

Leister D (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. Trends Genet 20:116–122

Logemann J, Schell J, Willmitzer L (1987) Improved method for the isolation of RNA from plant tissues. Anal Biochem 163:16–20

Luck JE, Lawrence GJ, Dodds PN, Shepherd KW, Ellis JG (2000) Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. Plant Cell 12:1367–1377

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

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003) Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. Plant Cell 15:809–834

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

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

Monforte AJ, Olive M, Gonzalo MJ, Alvarez JM, Dolcet-Sanjuan R, Arús P (2004) Identification of quantitative trait loci involved in fruit quality traits in melon (*Cucumis melo* L.). Theor Appl Genet 108:750–758

Oliver M, Garcia-Ma J, Cardús M, Pueyo N, López-Sesé AI, Arroyo M, Gómez-Paniagua H, Arús P, Vicente MC (2001) Construction of a reference linkage map for melon. Genome 44:836–845

Page RDM (1996) TREEVIEW: an application to display phylogenetic trees on personal computers. Comp Appl Biosci 12:357–358

Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BBH, Jones JDG (1997) Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf4/9* locus of tomato. Cell 91:821–832

Ray S, Anderson JM, Urmeev FI, Goodwin SB (2003) Rapid induction of a protein disulfide isomerase and defense-related genes in wheat in response to the hemibiotrophic fungal pathogen *Mycosphaerella graminicola*. Plant Mol Biol 53:701–714

Richly E, Kurth J, Leister D (2002) Mode of amplification and reorganization of resistance genes during recent *Arabidopsis thaliana* evolution. Mol Biol Evol 19:76–84

Sawant SV, Kiran K, Singh PK, Tuli R (2001) Sequence architecture downstream of the initiator codon enhances gene expression and protein stability in plants. Plant Physiol 126:1630–1636

Sharp PA (1981) Speculations on RNA splicing. Cell 23:643–646

Smith RF, Wiese BA, Wojzynski MK, Davison DB, Worley KC (1996) BCM Search Launcher—an integrated interface to molecular biology data base search and analysis services available on the World Wide Web. Genome Res 6:454–462

Soltis PS, Soltis DE, Chase MW (1999) Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. Nature 402:402–404

Staden R (1996) The Staden sequence analysis package. Mol Biotechnol 5:233–241

Temnykh S, DeClerck G, Lukashova A, Lipovich L., Cartinhour S, McCouch S. (2001) Computational and experimental analysis of microsatellites in rice *(Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential. Genome Res 11:1441–1452

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature 408:796–815

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Research 22:4673–4680

van Leeuwen H, Monfort A, Zhang HB, Puigdoménech P (2003) Identification and characterisation of a melon genomic region containing a resistance gene cluster from a constructed BAC library. Microcolinearity between *Cucumis melo* and *Arabidopsis thaliana*. Plant Mol Biol 51:703–718

Vander Hoorn RA, Kruijt M, Roth R, Brandwagt BF, Joosten MH, De Wit PJ (2001) Intragenic recombination generated two distinct *Cf* genes that mediate AVR9 recognition in the natural population of *Lycopersicon pimpinellifolium*. Proc Natl Acad Sci USA 98:10493–10498

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

Young ND (2000) The genetic architecture of resistance. Curr Opin Plant Biol 3:285–290

Zhang HB, Choi S, Woo SS, Li Z, Wing RA (1996) Construction and characterisation of two rice bacterial artificial chromosome libraries from the parents of a permanent recombinant inbred mapping population. Mol Breed 2:11–24

Zhou T, Wang Y, Chen JQ, Araki H, Jing Z, Jiang K, Shen J, Tian D (2004) Genome-wide identification of NBS genes in *japonica* rice reveals significant expansion of divergent non-TIR NBS-LRR genes. Mol Genet Genomics 271:402–415

Zhu H, Cannon SB, Young ND, Cook DR (2002) Phylogeny and genomic organization of the TIR and non-TIR NBS–LRR resistance gene family in *Medicago truncatula*. Mol Plant Microbe Interact 15:529–539