Isolation and characterization of NBS-encoding disease resistance gene analogs in watermelon against fusarium wilt

Anand C. Reddy[1], B. Lavanya[1], T. Tejaswi[1], E. Sreenivasa Rao[2] and D. C. Lakshmana Reddy[1,*]

[1]Division of Biotechnology, and [2]Division of Vegetable Crops, Indian Institute of Horticultural Research, Bengaluru 560 089, India

Fusarium wilt (*Fusarium oxysporum* f. sp. *niveum*) in watermelon is one of the deadliest diseases around the globe, and availability of disease-resistant varieties is moderate. Disease management utilizing resistance genes (R-genes)/resistance gene analogs (RGAs) has proven to be a promising and successful approach. In the present study, six watermelon RGAs were isolated from wild, fusarium wilt resistant genotype IIHR-82 (*Citrullus lanatus* var. *citroides*) using degenerate primers that identify nucleotide binding site-leucine-rich repeat (NBS–LRR) regions. Multiple sequence alignment of these RGAs identified the characteristic NBS–LRR motif, and BLASTp search revealed similarity of these RGAs with other pathogenesis-related proteins. Phylogeny and motif analysis revealed genetic diversity of RGAs within those isolated from watermelon and with other plant R-genes. The watermelon RGAs isolated in this study contained both TIR–NBS–LRR (TNL) and non-TIR–NBS–LRR (CNL) classes of R-genes. Protein secondary structure prediction of these watermelon RGAs revealed the composition of proteins, including α-helix, β-strand, disordered region and other template-related information. Watermelon RGAs identified in the present study will help in the development of RGA-based markers for resistance to fusarium wilt of watermelon.

Keywords: Disease management, fusarium wilt, resistance genes, watermelon.

WATERMELON (*Citrullus lanatus*) is an important commercial horticultural crop, with an annual worldwide production of 117 million tonnes[1]. The crop is prone to numerous devastating diseases that limit its cultivation. *Citrullus lanatus* var. *citroides* (*Clc*; IIHR-82), which is a related wild species of watermelon, has been reported to be resistant to several diseases such as fusarium wilt (*Fusarium oxysporum* f. sp. *niveum*; FON)[2], gummy stem blight (*Didymella bryoniae*)[3], root-knot nematode (*Meloidogyne incognita*)[4,5], powdery mildew[6,7], bacterial fruit blotch[8], anthracnose[9], Zucchini yellow mosaic virus (ZYMV)[10,11], watermelon bud necrosis virus (WBNV)[12], and papaya ringspot virus-watermelon strain (PRSV-W)[13]. Levi *et al.*[14,15] highlighted the low genetic diversity in cultivated watermelon and the need to use related species, especially *Citrullus lanatus* ssp. *citroides* to broaden the genetic base and introduce disease resistance properties. Fusarium wilt caused by FON is one of the oldest and economically factorable diseases of watermelon. Due to high survival rate in soil and evolution of highly unpredicted new races, fusarium wilt management in watermelon is complicated[16].

The infection mechanism of pathogens, including bacteria, virus, fungi, insects, oomycetes and nematodes differs considerably, but the R-gene products are surprisingly similar in these pathogens[17]. The vast majority of R-genes in plants encode nucleotide binding site-leucine-rich repeat (NBS–LRR) class of proteins. They have sequence homology with mammalian inflammatory and immune response proteins called nucleotide-binding oligomerization domain-LRR (NOD–LRR)[18]. The exact genetic mechanisms of NBS–LRR-based recognition and attack of a pathogen are not well understood[19]. The NBS domain contains ~300 amino acids and eight motifs arranged in a specific order. The domain is responsible for hydrolysis of nucleotide triphosphate (NTP) and signals transduction regulation via conformational changes in the protein structure[20]. The C-terminal region of LRRs normally contains 20–29 amino acid repeats that are implicated in the specific recognition of pathogen-derived virulence factors[21,22] and signalling phenomena[19,23]. The NBS–LRR genes are further classified into TIR–NBS–LRR (TNL) and non-TIR–NBS–LRR (CNL) classes, based on the TIR domain at the N-terminal. The CNL class is also called CC–NBS–LRR due to the presence of N-terminal coiled-coil (CC) domain[17]. Both TNL and CNL are involved in pathogen recognition, but differ in signalling pathways[24].

Currently, isolation and characterization of NBS-encoding RGAs are being done using degenerate primers for conserved domains of NBS with a high success rate[25]. Over 144 RGAs have been isolated from watermelon (*Citrullus lanatus* and *Citrullus colocynthis*) using

*For correspondence. (e-mail: lreddy@iihr.res.in)

Table 1. List of degenerate primers used in this study

S. no.	Primer	Primer sequence (5′–3′)	Conserved domains encoded	Reference
Primer 1	Ploop-F1	GGNGGNRTNGGNAAGACGAC	GG(I/M/V)GKTT	61
	GLPL-R1	GAGGGCTAAAGGAAGGCC	GLPLAL	62
Primer 2	Ploop-F2	TGSSRGGHWYRGGBAAAACTAC	(A/G/P/R)G(T/I/M/S/L)GKTT	63
	GLPL-R2	HRCWARAGGVARCCCTYBACA	GLPL(A/T)L	
Primer 3	CLRR-F	TTTTCGTGTTCAACGACG	LRR group	64
	CLRR-R	TAACGTCTATCGACTTCT		
Primer 4	RLRR-F	CGCAACCACTAGAGTAAC		
	RLRR-R	ACACTGGTCCATGAGGTT		
Primer 5	XLRR-F	CCGTTGGACAGGAAGGAG		
	XLRR-R	CCCATAGACCGGACTGTT		

degenerate primer-based PCR and *de novo* mining method[17,26,27]. The *I*, *I-2* and *I-3* resistant genes conferring resistance to *Fusarium oxysporum* f. sp. *lycopersici* (*Fol*) races 1, 2 and 3 respectively, were identified and these R-genes have similarities with NBS–LRR (*I* and *I-2*) (*or*) S-receptor-like kinase (SRLK; *I-3*) class of R-genes[28–32]. The other fusarium wilt dominant resistance gene called *resistance to Fusarium oxysporum* (*RFO1*) was isolated from *Arabidopsis thaliana*, and it confers resistance to a wide range of fusarium races[33].

In this study, we have used a fusarium wilt resistant accession species along with a susceptible variety to clone and characterize the RGAs, which can be of potential use in watermelon resistance breeding programmes.

Materials and methods

Fusarium wilt screening and genomic DNA isolation

In the present experiment, fusarium wilt inoculum was prepared according to Wechter *et al.*[34]. Watermelon wild species *C. lanatus* var. *citroides* (IIHR-82) resistant to fusarium wilt races 0, 1, and 2 was used along with fusarium wilt susceptible variety *C. lanatus* cv. Arka Manik for screening fusarium wilt resistance through artificial inoculation. Seeds were sown in pro-trays having 98 cells using sterilized coco peat. Only one seed was placed in each cell. Forty-two seeds of each variety were planted per test in two replications. One-week-old seedlings were used for artificial inoculation. Also, 5 ml of spore suspension at a concentration of 1×10^6 conidia/ml was used for drenching every seedling. The pro-trays were maintained in the temperature range 27–32°C and >70% relative humidity for good colonization and expression of symptoms. The plants were observed continuously for symptom development, and mortality was recorded daily. The total number of survival plants was recorded on a daily basis from the sixth day after inoculation for calculating the average daily survival percentage.

DNA was extracted from fresh young leaf tissue samples using modified CTAB (4%) method[35]. The quality and quantity of genomic DNA were determined using 0.8% TAE agarose gel and GeneQuant Pro spectrophotometer respectively.

Degenerate primers and PCR analysis

The NBS–LRR regions were amplified from IIHR-82 and *Citrullus lanatus* cv. Arka Manik using degenerate primers[36] (Table 1). These primers were designed by considering the conserved motifs of plant R-genes/RGAs. PCR amplification was performed in a 25 µl reaction volume. Each reaction mixture contained 1× PCR reaction buffer (Bangalore Genei, India), 2.5 mmol/l $MgCl_2$, 0.5 mmol/l dNTPs, 0.5 mmol/l degenerate primer pair, 40 ng template genomic DNA, and 3 units of *Taq* polymerase (Bangalore Genei, India). The PCR reaction was performed in Eppendorf mastercycler Pro S, the conditions were initial denaturation at 94°C for 4 min, followed by 35 cycles at 94°C for 30 s, 55°C for 30 s and 72°C for 60 s, and a final extension at 72°C for 10 min.

Cloning and sequencing

PCR products were run on a 1.5% (w/v) TAE agarose gel and stained with ethidium bromide. The expected bands (500–600 bp) were excised from the agarose gel and purified with NucleoSpin gel extraction kit (Macherey-Nagel (MN), Germany). The purified fragments were ligated into a pTZ57R/T vector (Thermo Scientific, USA), and ligated fragments were transformed into competent DH5α cells. The blue–white selective screening was carried out, and specific clones were grown in Luria broth overnight at 37°C with a shaker speed of 250 rpm. Plasmid DNAs were isolated using a plasmid isolation kit (Bioserve, India). The identity of each purified plasmid DNA was confirmed using restriction analysis and colony PCR.

Table 2. Identified similarities between shortlisted watermelon resistance gene analogs (RGAs) and GenBank accessions using BLASTp algorithm. Some of the shortlisted hits have high level of identity with existing R-genes of various plants

Watermelon RGAs	Similarity description	Plant species	Identity (%)	*E*-value
ClcRGA1	NADH dehydrogenase subunit 4 (APW82683)	*Citrullus lanatus* subsp. *vulgaris*	88	1e-84
	NADH-plastoquinone oxidoreductase subunit 4 (YP_004841835)	*Cucumis melo* subsp. *melo*	88	1e-82
ClcRGA2	Predicted: Solute carrier family 25 member 44 (XP_004141336)	*Cucumis sativus*	69	1e-28
	Solute carrier family 25 member 44 (XP_021903034)	*Carica papaya*	51	3e-22
ClcRGA3	Predicted: Low-quality protein: disease resistance protein RPP4-like (XP_016901470)	*Cucumis melo*	77	7e-45
	NBS-LRR resistance protein (AEV46173)	*Citrullus lanatus*	72	1e-43
ClcRGA10	Predicted: TMV resistance protein N-like (XP_016901592)	*Cucumis melo*	72	7e-69
	TIR-NBS-LRR-AAA + ATPase class resistance protein (ADI99934)	*Cucumis sativus*	73	3e-70

Representative plasmids with insert were selected for Sanger sequencing.

Domain identification and phylogeny

The vector region was deleted, and the remaining RGA-representing sequence containing conserved domains between P-loop and GLPL was considered for further analysis. The sequences of nucleotides were translated to the predicted amino acid sequence using the translation tool of Expasy. The sequence homology search was performed to identify homologous sequences through the BLASTp algorithm of the National Center for Biotechnology Information (NCBI), USA. The translated amino acid sequences were used for multiple sequence alignment and homology matrix with other highly characterized R-genes and previously reported full-length watermelon RGAs, using DNAMAN version 9 (Lynnon Biosoft, Canada). The RGAs were analysed for highly conserved domain/motif structures using MEME suite 4.12.0 (Multiple Expectation Maximization for Motif Elicitation)[37]. Phylogenetic and molecular evolutionary analyses of RGAs, identified in the present study, were performed along with previously reported watermelon RGAs and known R-genes using MEGA version 7.0 (Molecular Evolutionary Genetic Analysis)[38]. To construct a phylogenetic tree, the maximum likelihood method was used with 1000 bootstrap values. Hidden Markov model-based alignment method was used in Phyre 2 software for aligning and detecting secondary structure and modelling of watermelon R-genes/RGAs proteins[39]. The separate phylogenetic tree was also constructed with six isolated RGAs and fusarium wilt resistance genes of other crops to identify similarity and diversity within fusarium wilt resistance genes/RGAs.

Results

The NBS–LRR encoding regions (~500–600 bp) in *C. lanatus* var. *citroides* and *Citrullus lanatus* cv. Arka Manik was amplified with the help of specific degenerate primers. The amplicons were eluted using an elution and purification kit, and cloned into pTZ57R/T vector. All 250 clones were randomly selected for plasmid isolation, and 230 plasmids were confirmed as positive through restriction digestion and colony PCR. The positive plasmids were sequenced using Sanger's method. Among them, 30 had short-length sequences and were rejected. Similarity search in NCBI database using nucleotide BLAST (BLASTn) algorithm revealed that 75 clones had no hit, and were also excluded. The remaining 125 sequences were translated into polypeptides using the Expasy translate tool (https://web.expasy.org/translate/). However, nearly 86 of them contained frame shifts or stop codons, and/or did not contain any specific RGA representative domains. The remaining 39 translated sequences contained uninterrupted open reading frames (ORFs) and had RGAs representing domains with similarity to other plant R-genes named as ClcRGAs. Homology matrix was performed using the 39 sequences along with other highly characterized R-genes of different plants. The results identified replicates or repeated sequences. Six sequences with less than 67% homology among the 39 sequences were shortlisted for further analysis and deposited in GenBank database. Their accession numbers are: MF627703 (ClcRGA1), MF627704 (ClcRGA2), MF627705 (ClcRGA3), MF627706 (ClcRGA10), MF627707 (ClcRGA13) and MF627708 (ClcRGA30). Six protein sequences selected were used for BLASTp, and results showed a high similarity score with existing RGAs or R genes of different plants (Table 2).

Multiple sequence alignment and MEME analysis

Multiple sequence alignment was performed with all RGAs isolated from the present and previous studies, along with highly characterized R-genes, namely RPM1 (AGC12590), Gpa2 (AF195939), PRF (AAC49408) and M (AAB47618). Results demonstrated the presence of

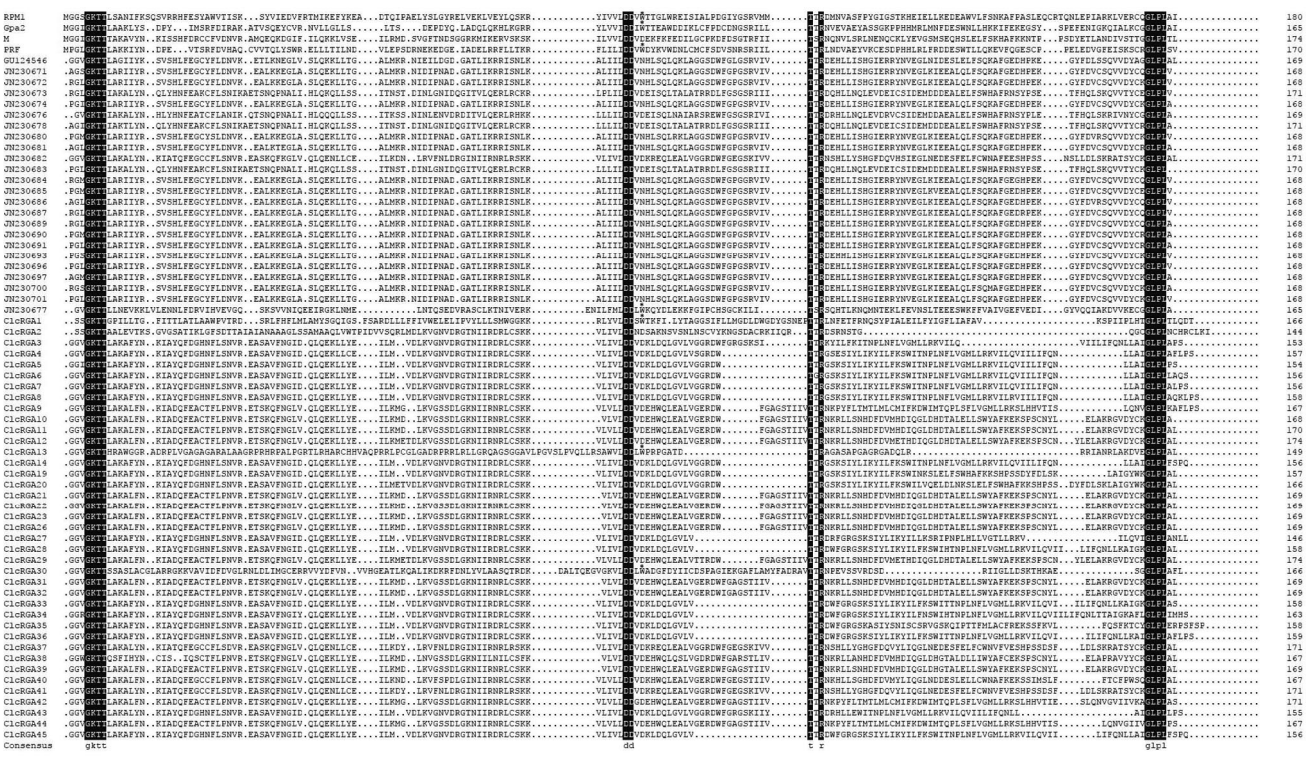

Figure 1. Multiple sequence alignment of watermelon resistance gene analogs of the present and previous studies along with known R-genes. *Non-TIR–NBS–LRR representative amino acid (W) in the RGA sequence.

different conserved domains in all the 39 watermelon RGAs, which had similarity with previously isolated watermelon RGAs and other known R-genes (Figure 1). Conserved domains such as P-loop, RNABS-A, kinase-2, RNBS-B, RNBS-C and GLPL were identified in watermelon RGAs. In the 39 RGAs, three sequences (ClcRGA1, ClcRGA13 and ClcRGA30) had tryptophan (W) after kinase-2 domain, explaining their similarity with other non-TIR–NBS–LRR class of RGAs. The remaining 36 sequences had aspartic acid (D) at the end of the kinase-2 motif, implying that these RGA sequences are related to the TIR–NBS–LRR type of RGAs.

Motif identification software MEME suite was used for the detection of conserved RGA domains in watermelon. The analysis was performed separately for TNL and CNL classes to understand diversity in watermelon RGAs. Presence of conserved motifs such as P-loop, RNBS-A, kinase-2, RNBS-B and HD-GLPL in watermelon RGAs confirmed their structural similarity with NBS–LRR class of R-genes (Figure 2). However, distribution of domains and their length differed in TNL and CNL of RGAs isolated from watermelon.

Phylogeny and diversity identification

The phylogenetic tree was primarily divided into two groups/clades: (I) TIR–NBS–LRR and (II) non-TIR–NBS–LRR class of *R*-genes (Figure 3). The TIR–NBS–LRR group was further classified into four subgroups: A, B, C, and D. Subgroup A consisted of TNL class of R-genes (ClcRGAs) of the present study and JN230682 from the previous study. Subgroups B and C consisted of the watermelon RGAs from previous studies, and subgroup D consisted of known characterized TIR-NBS-LRR gene *M*. From the present study ClcRGA1, ClcRGA13 and ClcRGA30 watermelon RGAs, and from previous studies DQ156558, DQ156559, DQ156560, GU124540, GU124542, GU124543 and JN230677 watermelon RGAs were grouped with highly characterized CNL sequences, such as RPM1, PRF and Gpa2. Phylogenetic analysis revealed that RGAs isolated in the present study differed from those reported previously.

The shortlisted six watermelon RGAs (ClcRGA1, ClcRGA2, ClcRGA3, ClcRGA10, ClcRGA13 and ClcRGA30) were used along with fusarium wilt resistance genes in other crops such as *I2* (accession no. ABB00396), *I2C-1* (AAB63274), *Fom-2* (AAS80152), *Fom-1* (AGH33855), *Fom-1* (AGH33848), 14-3-3 protein (AGZ13503), *RFO1* (AAY86486), R2R3-Myb transcription factor (AHZ33834), basic leucine zipper protein (AHI85805) and pathogenesis-related protein 10 families (AHG94646, AHG94647, AHG94650, AHG94651 and AHG94652) to construct a phylogenetic tree (Figure 4). The resulting phylogenetic tree was divided into two major groups: (i) R-genes mostly without NBS and LRR:

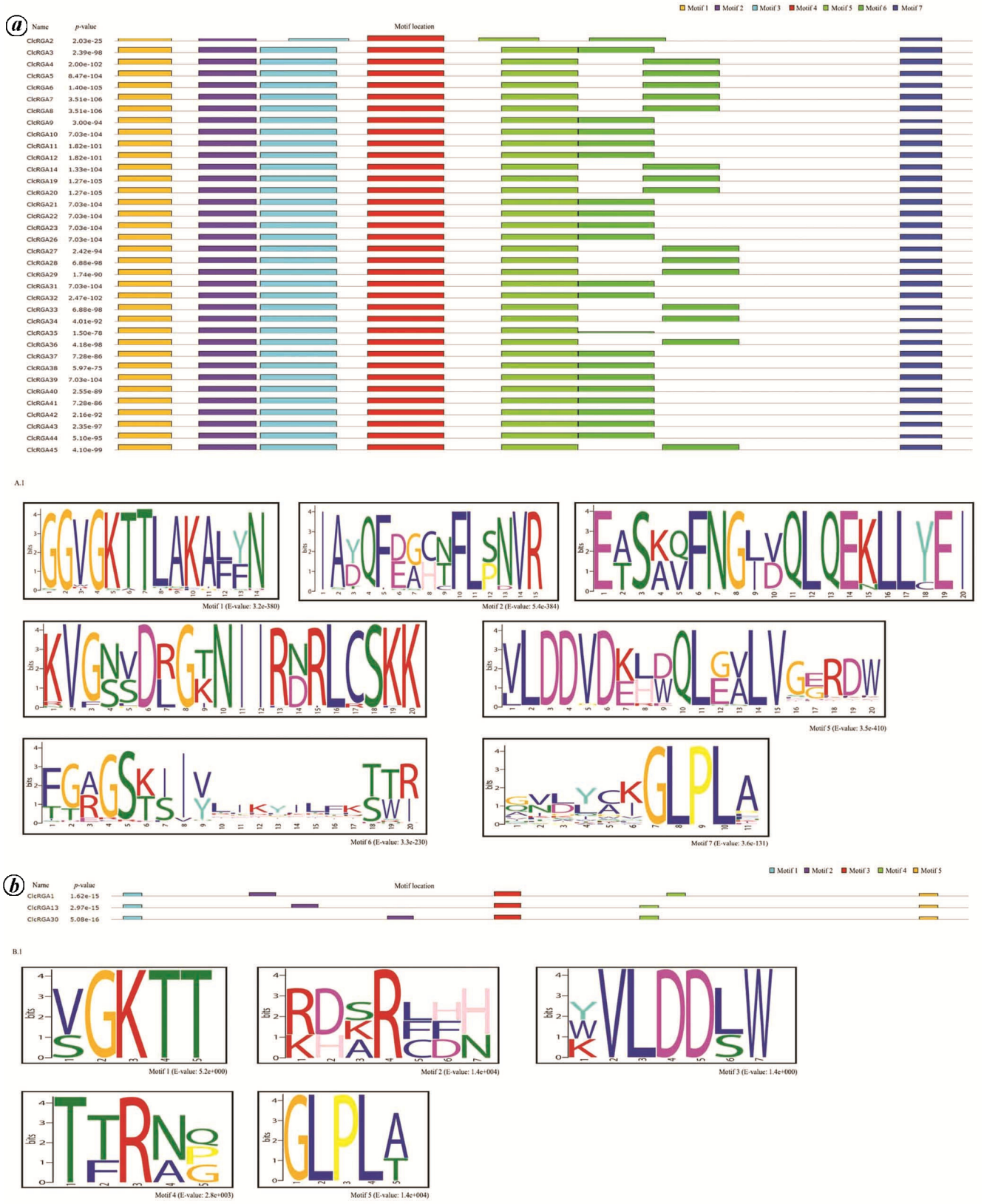

Figure 2. Conserved amino acid motifs identified within watermelon TNL and CNL classes of R-genes/RGAs separately, using MEME suite. Coloured boxes represent conserved motifs along the full length of watermelon RGAs. ***a***, Conserved motifs within the TNL class of R-genes/RGAs; A.1 Sequence signature of seven conserved motifs in TNL along with *E*-values. ***b***, Conserved motifs within the CNL class of R-genes/RGAs. B.1, Sequence signature of five conserved motifs in CNL along with *E*-values.

pathogenesis-related protein 10 genes, *RFO1*, basic leucine zipper gene, R2R3-Myb transcription factor and ClcRGA13. Also, ClcRGA13 was grouped with R2R3-Myb transcription factor; this may be due to less conserved regions within ClcRGA13 compared to other CNL classes of watermelon RGAs. (ii) NBS–LRR class of R-genes referring both TNL and CNL. The TNL subgroup contains ClcRGA2, ClcRGA3, ClcRGA10, *M* and *Fom-1* genes. The CNL subgroup contains ClcRGA1, ClcRGA30, Gpa2, Prf, RPM1, *Fom-2*, *I2* and *I2C-1*. The phylogenetic analysis confirms that the six isolated watermelon RGAs show similarity with fusarium wilt related R-genes of other crops.

Homology analysis and protein secondary structure prediction

The six shortlisted watermelon RGAs were compared with known, highly characterized R-genes such as RPM1, Gpa2, PRF and M, for understanding homology (Table 3). Homology of six watermelon RGAs with known R-genes ranged from 18.3% (ClcRGA13) to 47.3% (ClcRGA10). Homology among six watermelon RGAs varied from 19.4% to 66.7%, indicating a wide range of NBS-LRR diversity in the isolated watermelon RGAs.

Compared to primary structure (amino acid sequence), the secondary and tertiary structures of a protein are considered to be highly conserved during evolution, and hence are of relevance[40]. Six isolated watermelon RGAs were used to predict protein secondary structure and function, employing a web-based tool called Phyre 2 (ref. 39). It uses an efficient method called 'protein threading – a computational approach which finds compatibility of new protein sequence structure and functions using experimentally derived existing protein template information[41]. The present study revealed the composition of α-helix, β-strand, disordered sequence and other protein structure-related information along with template information (Table 4). The RGAs identified in this study consisted of 46%–85% α-helix, 0%–12% β-strand, and 8%–39% disordered sequence. Three sequences (ClcRGA1, ClcRGA10 and ClcRGA30) contained transmembrane helix regions that may have a role in specific interactions in the plane of lipid bi-layers[42].

Discussion

Watermelon is a major vegetable crop used globally. It is prone to huge yield losses due to diseases such as fusarium wilt, gummy stem blight, anthracnose, powdery mildew and downy mildew. Better understanding and knowledge of disease resistance through R-genes, which confer resistance to a wide range of pathogens, will be of help in watermelon breeding programmes. The NBS–LRR class of R-genes has been exclusively studied in various crops such as rice[43,44], *Arabidopsis*[45,46], cucumber[17], eggplant[36], tomato[47], melon[48] and pepper[49,50].

Various breeding programmes and their screening results have revealed that *C. lanatus* var. *citroides* is resistant to fusarium wilt races 0, 1, and 2 (ref. 26); gummy stem blight[3], root-knot nematode[4,5], and ZYMV and

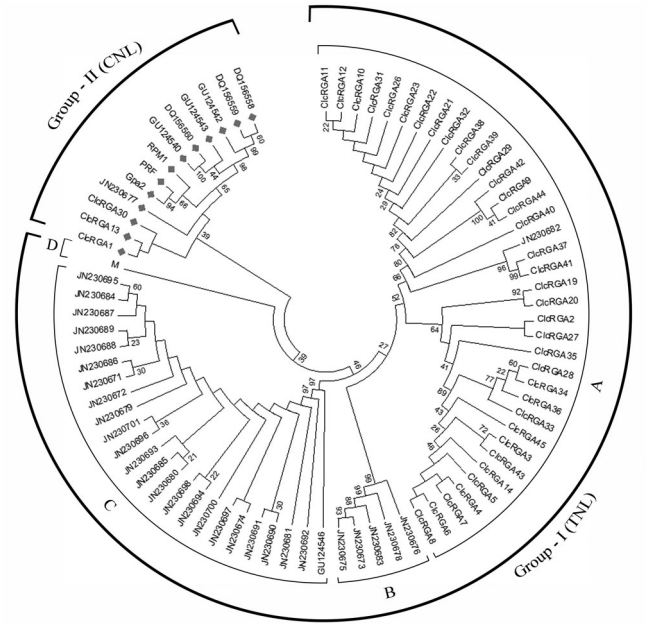

Figure 3. Phylogenetic tree of watermelon RGAs of the present and previous studies with known characterized R-genes using the maximum likelihood method. The bootstrap-based confidence values are mentioned at clades of the tree.

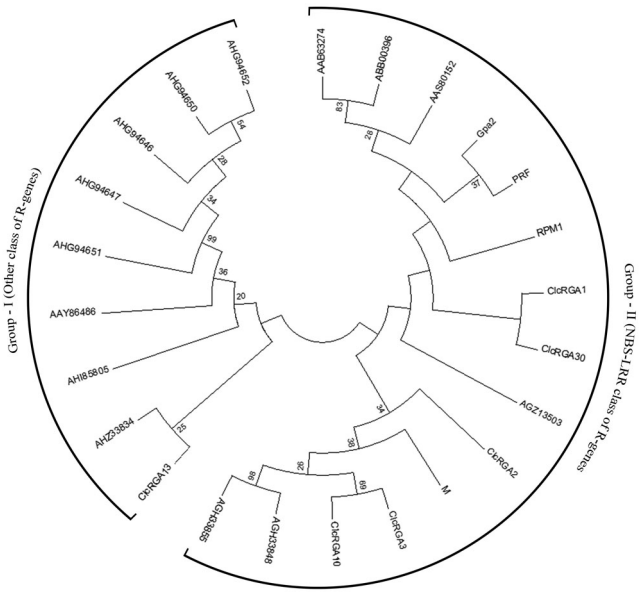

Figure 4. Phylogenetic tree of the shortlisted six watermelon RGAs along with fusarium wilt resistance genes in other crops using the maximum likelihood method. The bootstrap based confidence values are mentioned at clades of the tree.

Table 3. Homology matrix between watermelon RGAs and known characterized R-genes (%)

RPM1	100.0									
Gpa2	34.4	100.0								
M	25.8	29.0	100.0							
PRF	26.9	51.6	25.8	100.0						
ClcRGA1	26.9	22.6	22.6	22.6	100.0					
ClcRGA2	25.8	21.5	30.1	22.6	26.9	100.0				
ClcRGA3	25.8	26.9	45.2	31.2	22.6	44.1	100.0			
ClcRGA10	29.0	29.0	47.3	30.1	23.7	39.8	66.7	100.0		
ClcRGA13	23.7	25.8	18.3	20.4	21.5	20.4	21.5	20.4	100.0	
ClcRGA30	26.9	30.1	22.6	21.5	22.6	19.4	20.4	24.7	26.9	100.0

Table 4. Secondary structure composition and template information regarding watermelon RGAs and other fusarium wilt-related R-genes/RGAs

RGAs	α-Helix (%)	β-Strand (%)	Disordered (%)	TM helix (%)	Confidence	Identity (%)	Template information (fold/super-family/family)
ClcRGA1	85	0	12	60	100	16	Crystal structure of the membrane domain of respiratory complex I from *Escherichia coli* at 3.0 Å resolution
ClcRGA2	56	12	27	–	99.8	17	Structure of the *Drosophila* apoptosome
ClcRGA3	50	12	8	–	99.9	16	
ClcRGA10	56	8	15	10	99.9	18	Structure of a ced-4/ced-9 complex
ClcRGA13	46	0	39	–	36.6	24	Structure of the KdpFABC complex
ClcRGA30	47	12	16	10	100	74	The structure of the dimeric *E. coli* MinD-ATP complex

PRSV-W[11,13]. In the present study, the NBS–LRR class of R-genes was isolated from a fusarium wilt wild watermelon genotype *C. lanatus* var. *citroide,* and susceptible genotype *C. lanatus* cv. Arka Manik using the degenerate PCR approach. Using degenerate primers, a ~500 bp band was amplified, cloned and sequenced. Out of the 250 clones sequenced, 39 were identified as potential RGAs, and 6 of these 39 RGAs were shortlisted through sequence analysis for further characterization. Sequence analysis further revealed that the RGAs isolated have specific conserved motifs, suggesting their similarities with NBS–LRR class of R-genes.

Harris *et al.*[26] cloned 66 watermelon NBS–LRR class RGAs from fusarium wilt-resistant varieties, such as Calhoun Gray (resistance to race 0, 1), PI 595203 (resistance to race 0, 1, 2), and PI 296341 (multi-viral resistant), using degenerate primer-based PCR method. Wan *et al.*[17] identified 37 RGAs (two pseudogenes) from *C. lanatus* and *Citrullus colocynthis* using degenerate PCR and data mining. Recently, using genome sequencing and *de novo* methods two separate studies have identified a total of 99 watermelon RGAs (24 pseudogenes), which are related to both TNL and CNL classes of R-genes[27,51]. Recently, Lambel *et al.*[52] identified a major quantitative trait locus (QTL) associated with fusarium wilt race 1 in watermelon chromosome-1 through selective genotyping and GBS method. Apart from watermelon, fusarium wilt resistance source was identified and linked markers were reported in other crops such as melon and pea. Joobeur *et al.*[53] identified a single resistance gene for fusarium wilt resistance locus *Fom-2* in melon and is related to the NBS–LRR class of R-genes. In melon, one tightly linked AFLP marker and three closely mapped AFLP markers (converted to STS and CAPS) near *Fom-1* gene were identified and linkage map constructed[54]. In peas, *Fw* gene conferring resistance to fusarium wilt race-1 was mapped to linkage group III (ref. 55). The RGAs identified in the present study complement those reported in previous studies.

Results of multiple sequence alignment, motif analysis, homology matrix and phylogenetic analysis in watermelon RGAs with known R-genes suggest wide genetic diversity within the isolated RGAs. Multiple sequence alignment of RGAs revealed significant homology with previously reported watermelon RGAs and known R-genes. These results suggest that RGAs isolated may function as R-genes. When all 39 sequences were used for MEME motif analysis, some of the motifs highlighted in the multiple sequence analysis, were not detected, suggesting that they are poorly conserved in the protein sequence. These results are consistent with a previous study[17]. TNL sequences are homogeneous and, unlike CNL counterparts, mostly present in dicots. For this reason, MEME analysis was performed separately for TNL and CNL classes of watermelon RGAs.

Among 39 RGAs, 3 (ClcRGA1, ClcRGA13 and ClcRGA30) had conserved residue tryptophan at the end of the kinase-2 domain, and hence belong to CNL class of R-genes. The remaining 36 RGAs had highly conserved aspartic acid residue at the end of the kinase-2 motif, implying they fit into the TNL class. Aspirate or tryptophan residue at the end of kinase-2 domain suggests class

of NBS–LRR with 95% accuracy[24]. Multiple sequence alignment and phylogenetic analysis also corroborated this finding. As observed in the present study, the existence of both TNL and CNL is reported in eggplant[36], grapevine[56], peppers[49,50] and cucumber[57]. According to previous reports, monocots have only the CNL class of R-genes, while dicots have both TNL and CNL classes of R-genes[36,58]. The present study also confirms that watermelon has both TNL and CNL classes of R-genes like other dicots.

In the present study, the ratio of TNL to CNL classes of R-genes was observed to be 1 : 1, whereas Harris *et al.*[26] observed a ratio of 5 : 3 (TNL : CNL) in watermelon. However, these studies may not reflect the true ratio of TNL to CNL genes in the watermelon genome. For example, TNL : CNL ratio in the *Arabidopsis* genome was reported to be 2 : 1. However, using different degenerate primer combinations and PCR strategies, the ratio varied from 1 : 4 to 1 : 6 in different studies[56,59,60]. This can be explained by the fact that different studies use a variety of degenerate primer combinations and PCR conditions.

BLASTp results revealed identity of the six isolated RGAs with existing R-genes of other crops (Table 2). ClcRGA3 and ClcRGA10 showed identity with disease-resistant RPP4-like protein and TMV resistance N-like protein respectively. These results indicate that the watermelon RGAs from the present study may have a related role in disease resistance. The phylogenetic tree confirms the existence of both classes of NBS–LRR RGAs in watermelon.

The plant protein secondary structure-related information in *R*-genes/RGAs is scarce and not available for watermelon. As limited, experimentally determined protein tertiary (3D) structures are available, it is difficult to predict the exact structure or function of any protein of interest[39]. Hence, *in silico* protein secondary structure prediction and modelling were performed using Phyre 2, with the shortlisted watermelon RGAs. The results revealed similarity of RGAs with aligned templates of other characterized protein sequences. The composition of α-helix, β-strand and disordered sequences varied in watermelon RGAs. The results revealed similarity of RGAs with aligned templates of other characterized protein sequences. The composition of α-helix, β-strand and disordered sequences varied in watermelon RGAs. This reveals similarity with existing protein sequences and less problematic nature of isolated RGAs upon protein folding. Interestingly, transmembrane helix was observed in some of the RGAs such as ClcRGA1, ClcRGA10 and ClcRGA30. ClcRGA1 and ClcRGA13 contain α-helix and lack β-strand, and have similarity with respiratory complex and KdpFABC complex respectively. ClcRGA2 and ClcRGA3 have both α-helix and β-strand regions, and have similarity with *Drosophila* apoptosome protein structure. ClcRGA10 and ClcRGA30 have similarity with the structure of a ced-4/ced-9 complex and dimeric *Escherichia coli* MinD-ATP complex respectively. Additionally, most of the RGAs have a lesser composition of the disordered region (<50%).

Conclusion

In conclusion, degenerate PCR amplification-based method was used for identifying watermelon RGAs. Six RGAs belonging to NBS–LRR class of R-genes were characterized. The diversity of watermelon RGAs isolated was confirmed by phylogeny and multiple sequence alignment. These RGAs belong to both TNL and CNL classes of R-genes. The six watermelon RGAs developed in this study can be used to characterize resistance genes in watermelon and other cucurbitaceous crops. The RGAs reported can be used in the development of a molecular marker for plant disease resistance. Our findings provide a pathway to identify candidate R-genes related to watermelon diseases, and will help evaluate the NBS–LRR class of R-genes in watermelon.

Conflict of interest: All authors declare that they have no conflict of interest.

1. FAO, Food and Agriculture Organization of the United Nations – Statistics Division, 2016; http://www.fao.org/faostat/en/#data/QC (accessed on 23 March 2018).
2. Martyn, R. D. and Netzer, D., Resistance to races 0, 1 and 2 of Fusarium wilt of watermelon in *Citrullus* sp. PI-296341-FR. *HortScience*, 1991, **26**(4), 429–432.
3. Gusmini, G., Song, R. and Wehner, T. C., New sources of resistance to gummy stem blight in watermelon. *Crop Sci.*, 2005, **45**(2), 582–588.
4. Thies, J. A. and Levi, A., Resistance of watermelon germplasm to the peanut root-knot nematode. *HortScience*, 2003, **38**(7), 1417–1421.
5. Thies, J. A. and Levi, A., Characterization of watermelon (*Citrullus lanatus* var. *citroides*) germplasm for resistance to root-knot nematodes. *HortScience*, 2007, **42**(7), 1530–1533.
6. Thomas, C. E., Levi, A. and Caniglia, E., Evaluation of US plant introductions of watermelon for resistance to powdery mildew. *HortScience*, 2005, **40**(1), 154–156.
7. Davis, A. R., Levi, A., Wehner, T. and Pitrat, M., PI 525088-PMR, a melon race 1 powdery mildew-resistant watermelon line. *HortScience*, 2006, **41**(7), 1527–1528.
8. Hopkins, D. L. and Thompson, C. M., Evaluation of *Citrullus* sp. germ plasm for resistance to *Acidovorax avenae* subsp. *citrulli*. *Plant Dis.*, 2002, **86**, 61–64.
9. Sowell Jr, G., Rhodes, B. B. and Norton, J. D., New sources of resistance to watermelon anthracnose. *J. Am. Soc. Hortic. Sci.*, 1980, **105**, 197–199.
10. Provvidenti, R., Inheritance of resistance to the Florida strain of zucchini yellow mosaic virus in watermelon. *HortScience*, 1991, **26**(4), 407–408.
11. Harris, K. R., Ling, K. S., Wechter, W. P. and Levi, A., Identification and utility of markers linked to the zucchini yellow mosaic virus resistance gene in watermelon. *J. Am. Soc. Hortic. Sci.*, 2009, **134**(5), 529–534.

12. Swamy, K. M. *et al.*, Development and evaluation of an interspecific population involving *Citrullus lanatus* var. *citroides* to develop pre-bred lines for resistance to WBNV in watermelon. In VIROCON, Indian Institute of Horticultural Research, Bengaluru, 2016.
13. Guner, N. and Wehner, T. C., Overview of potyvirus resistance in watermelon. In Proceedings of the IX EUCARPIA Meeting Cucurbitaceae, l'institut National de la Recherche Agronom-ique, Avignon, France, 2008.
14. Levi, A., Wechter, P., Massey, L., Carter, L. and Hopkins, D., An extended genetic linkage map for watermelon based on a testcross and a BC2F2 population. *Am. J. Plant Sci.*, 2011, **2**(2), 93.
15. Levi, A. *et al.*, High frequency oligonucleotides: targeting active gene (HFO-TAG) markers revealed wide genetic diversity among *Citrullus* spp. accessions useful for enhancing disease or pest resistance in watermelon cultivars. *Genet. Resour. Crop Evol.*, 2013, **60**(2), 427–440.
16. Egel, D. S. and Martyn, R. D., Fusarium wilt of watermelon and other cucurbits. *The Plant Health Instructor*, 2007; doi:10.1094/PHI-I-2007-0122-01 (updated 2013).
17. Wan, H., Yuan, W., Bo, K., Shen, J., Pang, X. and Chen, J., Genome-wide analysis of NBS-encoding disease resistance genes in *Cucumis sativus* and phylogenetic study of NBS-encoding genes in Cucurbitaceae crops. *BMC Genomics*, 2013, **14**(1), 109.
18. Inohara, N., Chamaillard, M., McDonald, C. and Nunez, G., NOD–LRR proteins: role in host–microbial interactions and inflammatory disease. *Ann. Rev. Biochem.*, 2005, **74**, 355–383.
19. DeYoung, B. J. and Innes, R. W., Plant NBS-LRR proteins in pathogen sensing and host defense. *Nature Immunol.*, 2006, **7**, 1243–1249.
20. Martin, G. B., Bogdanove, A. J. and Sessa, G., Understanding the function of plant disease resistance proteins. *Annu. Rev. Plant Biol.*, 2003, **54**, 23–61.
21. Kobe, B. and Kajava, A. V., The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol.*, 2001, **11**(6), 725–732.
22. Ellis, J. G., Lawrence, G. J., Luck, J. E. and Dodds, P. N., Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. *Plant Cell*, 1999, **11**, 495–506.
23. Luck, J. E., Lawrence, G. J., Dodds, P. N., Shepherd, K. W. and Ellis, J. G., Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. *Plant Cell*, 2000, **12**, 1367–1377.
24. Meyers, B. C., Dickerman, A. W., Michelmore, R. W., Sivaramakrishnan, S., Sobral, B. W. and Young, N. D., Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J.*, 1999, **20**(3), 317–332.
25. Sekhwal, M. K., Li, P., Lam, I., Wang, X., Cloutier, S. and You, F. M., Disease resistance gene analogs (RGAs) in plants. *Int. J. Mol. Sci.*, 2015, **16**(8), 19248–19290.
26. Harris, K. R., Wechter, W. P. and Levi, A., Isolation, sequence analysis, and linkage mapping of nucleotide binding site–leucine-rich repeat disease resistance gene analogs in watermelon. *J. Am. Soc. Hortic. Sci.*, 2009, **134**(6), 649–657.
27. Lin, X., Zhang, Y., Kuang, H. and Chen, J., Frequent loss of lineages and deficient duplications accounted for low copy number of disease resistance genes in Cucurbitaceae. *BMC Genomics*, 2013, **14**, 335.
28. Ori, N. *et al.*, The I2C family from the wilt disease resistance locus I2 belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *Plant Cell*, 1997, **9**(4), 521–532.
29. Simons, G. *et al.*, Dissection of the Fusarium I2 gene cluster in tomato reveals six homologs and one active gene copy. *Plant Cell*, 1998, **10**(6), 1055–1068.
30. McGrath, D. J., Gillespie, D. and Vawdrey, L., *Inheritance* of resistance to *Fusarium oxysporum* f. sp. *lycopersici* races 2 and 3 in *Lycopersicon pennellii*. *Aust. J. Agric. Res.*, 1997, **38**, 729–733.
31. Scott, J. W. and Jones, J. P., Monogenic resistance in tomato to *Fusarium oxysporum* f. sp. *lycopersici* race 3. *Euphytica*, 1989, **40**, 49–53.
32. Catanzariti, A. M., Lim, G. T. and Jones, D. A., The tomato I-3 gene: a novel gene for resistance to *Fusarium* wilt disease. *New Phytol.*, 2015, **207**(1), 106–118.
33. Diener, A. C. and Ausubel, F. M., Resistance to *Fusarium oxysporum* 1, a dominant *Arabidopsis* disease-resistance gene, is not race specific. *Genetics*, 2005, **171**(1), 305–321.
34. Wechter, W. P., Kousik, C., McMillan, M. and Levi, A., Identification of resistance to *Fusarium oxysporum* f. sp. *niveum* race 2 in *Citrullus lanatus* var. *citroides* plant introductions. *HortScience*, 2012, **47**(3), 334–338.
35. Murray, M. G. and Thompson, W. R., Rapid isolation of high molecular weight plant DNA. *Nucl. Acids Res.*, 1980, **8**, 4321–4325.
36. Reddy, A. C., Venkat, S., Singh, T. H., Aswath, C., Reddy, K. M. and Reddy, D. L., Isolation, characterization and evolution of NBS–LRR encoding disease-resistance gene analogs in eggplant against bacterial wilt. *Eur. J. Plant Pathol.*, 2015, **143**(3), 417–426.
37. Bailey, T. L. *et al.*, MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, 2009, **37**, W202–W208.
38. Kumar, S., Stecher, G. and Tamura, K., MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.*, 2016, **33**(7), 1870–1874.
39. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. and Sternberg, M. J., The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Proto.*, 2015, **10**(6), 845–858.
40. Koonin, E. V., Wolf, Y. I. and Karev, G. P., The structure of the protein universe and genome evolution. *Nature*, 2002, **420**(6912), 218.
41. Shao, M., Wang, S. and Wang, C., Incorporating *ab initio* energy into threading approaches for protein structure prediction. *BMC Bioinfor.* (*Suppl 1*), 2011, **12**, S54.
42. Zviling, M., Kochva, U. and Arkin, I. T., How important are transmembrane helices of bitopic membrane proteins? *Biochim. Biophys. Acta (BBA)-Biomembr.*, 2007, **1768**(3), 387–392.
43. Bryan, G. T. *et al.*, A single amino acid difference distinguishes resistant and susceptible alleles of the rice blast resistance gene *Pi-ta*. *Plant Cell*, 2000, **12**(11), 2033–2045.
44. Okuyama, Y. *et al.*, A multifaceted genomics approach allows the isolation of the rice Pia-blast resistance gene consisting of two adjacent NBS–LRR protein genes. *Plant J.*, 2011, **66**(3), 467–479.
45. Tan, X. *et al.*, Global expression analysis of nucleotide binding site–leucine-rich repeat-encoding and related genes in *Arabidopsis*. *BMC Plant Biol.*, 2007, **7(1)**, 56.
46. Borhan, M. H., Gunn, N., Cooper, A., Gulden, S., Tör, M., Rimmer, S. R. and Holub, E. B., WRR4 encodes a TIR–NB–LRR protein that confers broad-spectrum white rust resistance in *Arabidopsis thaliana* to four physiological races of *Albugo candida*. *Mol. Plant–Microb. Interact.*, 2008, **21**(6), 757–768.
47. Zhang, C. *et al.*, The *Ph-3* gene from *Solanum pimpinellifolium* encodes CC–NBS–LRR protein conferring resistance to *Phytophthora infestans*. *Theoret. Appl. Genet.*, 2014, **127**(6), 1353–1364.
48. Daryono, B. S., Wakui, K. and Natsuaki, K. T., Linkage analysis and mapping of SCAR markers linked to CMV-B2 resistance gene in melon. *SABRAO J. Breed. Genet.*, 2010, **42**(1), 35–45.
49. Wan, H. *et al.*, Analysis of TIR- and non-TIR–NBS–LRR disease resistance gene analogous in pepper: characterization, genetic variation, functional divergence and expression patterns. *BMC Genomics*, 2012, **13**(1), 502.

50. Naresh, P., Reddy, M. K., Reddy, A. C., Lavanya, B., Reddy, D. L. and Reddy, K. M., Isolation, characterization and genetic diversity of NBS–LRR class disease-resistant gene analogs in multiple virus resistant line of chilli (*Capsicum annuum* L.). *3 Biotech*, 2017, **7**(2), 114.
51. Guo, S. *et al.*, The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nature Genet.*, 2013, **45**(1), 51.
52. Lambel, S. *et al.*, A major QTL associated with *Fusarium oxysporum* race 1 resistance identified in genetic populations derived from closely related watermelon lines using selective genotyping and genotyping-by-sequencing for SNP discovery. *Theoret. Appl. Genet.*, 2014, **127**(10), 2105–2115.
53. Joobeur, T., King, J. J., Nolin, S. J., Thomas, C. E. and Dean, R. A., The fusarium wilt resistance locus *Fom-2* of melon contains a single resistance gene with complex features. *Plant J.*, 2004, **39**(3), 283–297.
54. Tezuka, T., Waki, K., Yashiro, K., Kuzuya, M., Ishikawa, T., Takatsu, Y. and Miyagi, M., Construction of a linkage map and identification of DNA markers linked to *Fom-1*, a gene conferring resistance to *Fusarium oxysporum* f. sp. melonis race 2 in melon. *Euphytica*, 2009, **168**(2), 177.
55. Okubara, P. A., Keller, K. E., McClendon, M. T., Inglis, D. A., McPhee, K. E. and Coyne, C. J., Y15_999Fw, a dominant SCAR marker linked to the Fusarium wilt race 1 (Fw) resistance gene in pea. *Pisum Genet.*, 2005, **37**, 30–33.
56. Donald, T. M., Pellerone, F., Adam-Blondon, A. F., Bouquet, A., Thomas, M. R. and Dry, I. B., Identification of resistance gene analogs linked to a powdery mildew resistance locus in grapevine. *Theoret. Appl. Genet.*, 2002, **104**, 610–618.
57. Wan, H., Zhao, Z., Malik, A. A., Qian, C. and Chen, J., Identification and characterization of potential NBS-encoding resistance genes and induction kinetics of a putative candidate gene associated with downy mildew resistance in *Cucumis*. *BMC Plant Biol.*, 2010, **10**, 186.
58. Joshi, R. K., Mohanty, S., Subudhi, E. and Nayak, S., Isolation and characterization of NBS–LRR-resistance gene candidates in turmeric (*Curcuma longa* cv. *surama*). *Genet. Mol. Res.*, 2010, **9**, 1796–1806.
59. Aarts, N., Metz, M., Holub, E., Staskawicz, B. J., Daniels, M. J. and Parker, J. E., Different requirements for EDS1 and NDR1 by disease resistance genes define at least two *R* genemediated signaling pathways in *Arabidopsis*. *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 10306–10311.
60. Speulman, E., Bouchez, D., Holub, E. B. and Beynon, J. L., Disease resistance gene homologs correlate with disease resistance loci of *Arabidopsis thaliana*. *Plant J.*, 1998, **14**(4), 467–474.
61. Noir, S., Combes, M.-C., Anthony, F. and Lashermes, P., Origin, diversity, and evolution of NBS-type disease-resistance gene homologues in coffee trees (*Coffea* L.). *Mol. Genet. Genom.*, 2001, **265**, 654–662.
62. Deng, Z. *et al.*, Cloning and characterization of NBS–LRR class resistance gene candidate sequences in citrus. *Theoret. Appl. Genet.*, 2000, **101**, 814–822.
63. Zhang, L., Chen, R., Zhang, J., Ouyang, B., Xiao, J., Li, H. and Ye, Z., Cloning and analysis of resistance gene analogs from pepper (*Capsicum annuum* L.). *Sci. Agric. Sin.*, 2008, **41**(1), 169–175.
64. Chen, X. M., Line, R. F. and Leung, H., Genome scanning for resistance-gene analogs in rice, barley, and wheat by high-resolution electrophoresis. *Theoret. Appl. Genet.*, 1998, **97**(3), 345–355.

ACKNOWLEDGEMENTS. We thank the Director, Indian Institute of Horticultural Research Bengaluru for providing research facilities and funding. We also thank Dr Leela Sahijram (Principal Scientist, ICAR-IIHR) for critical analysis and editing of the manuscript.

Received 22 June 2018; revised accepted 22 May 2019

doi: 10.18520/cs/v117/i4/617-626