Theor Appl Genet (2005) 110: 1074–1085
DOI 10.1007/s00122-005-1928-5

ORIGINAL PAPER

Doug J. Hinchliffe · Yingzhi Lu · Carol Potenza
Champa Segupta-Gopalan · Roy G. Cantrell
Jinfa Zhang

Resistance gene analogue markers are mapped to homeologous chromosomes in cultivated tetraploid cotton

Received: 13 October 2004 / Accepted: 6 January 2005 / Published online: 22 February 2005

Abstract Degenerate primers designed from conserved motifs of known plant resistance gene products were used to amplify genomic DNA sequences from the root-knot nematode (*Meloidogyne incognita*) resistance genetic source, Upland cotton (*Gossypium hirsutum*) cultivar Auburn 634 RNR. A total of 165 clones were isolated, and sequence analysis revealed 57 of the clones to be novel nucleotide sequences, many containing the resistance (R)-protein nucleotide-binding site motif. A cluster analysis was performed with resistance gene analogue (RGA) nucleotide sequences isolated in this study, in addition to 99 cotton RGA nucleotide sequences already deposited in GenBank, to generate a phylogenetic tree of cotton R genes. The cotton RGA nucleotide sequences were arranged into 11 groups and 56 sub-groups, based on genetic distances. Multiple sequence alignments were performed on the RGA sequences of each sub-group, and either the consensus sequences or individual RGA sequences were used to design 61 RGA-sequence-tagged site primers. A recombinant inbred line (RIL) population of cultivated tetraploid cotton was genotyped using RGA-specific primers that amplified polymorphic fragments between the two RIL parents. Nine RGA markers were mapped to homeologous chromosomes 12 and 26, based on linkage to existing markers that are located on these chromosomes.

Communicated by C. Möllers

D. J. Hinchliffe · Y. Lu · C. Potenza
C. Segupta-Gopalan · J. Zhang (✉)
Department of Agronomy and Horticulture,
New Mexico State University, Las Cruces,
NM 88003, USA
E-mail: jinzhang@nmsu.edu

R. G. Cantrell
Cotton Incorporated, 6399 Weston Parkway,
Cary, NC 27513, USA

Introduction

The cultivated tetraploids of the genus *Gossypium* represent one of the most economically important crops in the world. Based on worldwide estimates for 2002, cotton was harvested on 305,000 km^2 of farmland, producing 19.2 billion kg of raw cotton (US Department of Agriculture 2003). In the United States, the predominant cotton crop is Upland cotton (*G. hirsutum* L.), with cultivars of the extra-long stable Pima cotton (*G. barbadense* L.) grown predominantly in the southwestern region. As with all cultivated plant species, biotic stresses represent a major challenge to the farmer and can greatly impact seasonal yields. In the United States in 2003, the total estimated cotton loss due to disease was 2,856,015 bales, which equaled a total cotton yield loss of 12.16% (Blasingame and Patel 2004). Plants have evolved a complex immune system that allows recognition of specific pathogens by a specialized group of gene products known as resistance (R) proteins. Plant R proteins confer resistance to a wide variety of pathogenic organisms including insects, fungi, and bacteria. Identification of DNA markers and/or gene(s) involved in resistance to insects and other pathogens is of great interest in the cotton farming industry. Isolation and sequencing of cotton R-gene-related sequences could provide valuable information leading to the development of DNA markers linked to a variety of specific disease resistance phenotypes.

Cellular responses involved in plant disease resistance are represented by a large family of R genes whose gene products are characterized by several key features including nucleotide-binding site (NBS) domains, leucine-rich repeat (LRR) regions, transmembrane (TM) regions, and serine/threonine kinase catalytic domains (reviewed in Martin et al. 2003). Plant resistance gene analogues (RGAs) have been isolated from numerous plant species, using degenerate primers designed mostly from the predominant NBS and LRR regions of well-characterized R-gene

products (Leister et al. 1996; Yu et al. 1996; Chen et al. 1998; Collins et al. 1998; Mago et al. 1999; Madsen et al. 2003; Tan et al. 2003). The use of degenerate R-gene primers has proven very effective in identifying R genes from very diverse species of plants. In the spectrum of applied research, crop scientists have benefited greatly from identification and mapping of R genes that are linked to pathogen and pest resistance phenotypes. Identification of new sequence-specific R-gene markers that are associated with resistance phenotypes facilitates marker-assisted selection (MAS) and allows for introgression of resistance genes into susceptible elite cultivars. In this instance, the molecular mechanisms that confer the resistance phenotype may remain unknown to the plant breeder, but through the use of DNA markers, the loci containing the resistance genes of choice can be tracked through generations of selection. The breeding and selection of *Theobroma cacao* has recently benefited from the identification of RGA DNA markers designed from the NBS and LRR domains of known resistance genes (Kuhn et al. 2003) and the DNA-binding domain and other conserved motifs of the WRKY family of plant transcriptional regulators (Borrone et al. 2004). DNA markers that allow MAS of plants resistant to specific pathogens is of particular importance to a crop such as cacao, due to long and costly yield trials (5–8 years) and self-incompatibility of many cultivars (Kuhn et al. 2003). Cultivated cotton is classified as an indeterminate perennial; however, in the United States it is produced as an annual crop based on region-specific growing seasons. Cotton breeding programs in the United States focus on improved resistance to a variety of pathogens, integrating pathogen resistance phenotypes into high-yielding, high-fiber quality cultivars. Conventional cotton breeding selections rely mainly on bioassays such as severity of disease symptoms (e.g., galling indices), and on the ability of the breeder to identify desirable traits for generation advancement. These techniques can be extremely subjective and in the case of bioassays, can lead to false selections due to environmental factors such as non-uniform distribution of pathogens in the field. Clearly, identification of R-gene-linked DNA markers and/or genes that confer natural resistance to pests would be greatly beneficial to the cotton farmer in terms of overhead expenditures.

In this study, we report:

1. Isolation of novel RGA nucleotide sequences from Upland cotton, using degenerate primers designed from conserved motifs of known resistance gene products.
2. Cluster analysis and grouping by genetic distance of 154 RGA nucleotide sequences from *G. hirsutum* and *G. barbadense*.
3. Design of RGA consensus–sequence-specific primers representing RGA sub-groups from the cluster analysis.
4. Mapping of nine RGA sub-groups to cotton homeologous chromosomes 12 and 26, using a recombinant inbred line (RIL) population from an interspecific hybrid between Upland cotton and Pima cotton.
5. Amino acid sequence conservation among R-protein sequences represented by RGA-sequence-tagged site (STS) markers mapped to homeologous chromosomes 12 and 26.

Materials and methods

Genomic DNA extraction and PCR amplification of RGAs

For comparison purposes, genomic DNA from the near-isogenic lines (NILs), root-knot nematode (RKN)-resistant Auburn 634 RNR (Shepherd 1974), and its susceptible recurrent parent, Auburn 56, were isolated from cotton leaf tissues by first grinding young leaf tissues in liquid nitrogen with a mortar and pestle and then processing powderized tissue with the DNeasy Plant Mini Kit (Qiagen, Valencia, Calif., USA). Genomic DNA was quantitated by UV absorption at 260 nm and checked for integrity by resolving on a 0.8% agarose gel. Cotton RGAs were isolated using eight degenerate primer pairs (Table 1) as reported by Leister et al. (1996), Yu et al. (1996), Feuillet et al. (1997), and Chen et al. (1998). The genomic DNA isolated from the two NILs was used as templates in the PCR reactions. PCR reactions were performed in a PE Applied Biosystems GeneAmp PCR System 9700 (Applied Biosystems, Foster City, Calif., USA) in 30-µl volumes with the following concentrations: 4 ng/µl genomic DNA template, 0.4 mM dNTPs, 1.5 mM $MgCl_2$, 2 µM of each primer, and 0.1 U/µl of AmpliTaq Gold DNA polymerase (Applied Biosystems). Thermal cycling conditions followed Chen et al. (1998), with modifications in that the initial denaturation of 10 min at 94°C was followed by 45 cycles at 94°C for 1 min, 42°C for 1 min, and 72°C for 2 min. A final extension was performed at 72°C for 5 min. The annealing temperature of 42°C was chosen to optimize the visual PCR amplification products (Noir et al. 2001; Tan et al. 2003).

Isolation and cloning of RGAs

RGA amplification products from the two NILs were resolved on a 1% agarose gel and visualized by staining with ethidium bromide. However, only RGA bands from Auburn 634 RNR were excised from the agarose gel and purified with the QIAquick Gel Extraction Kit (Qiagen). A total of 23 bands were isolated. Bands that were unsuccessfully cloned from the first purification were reamplified using the same degenerate primer pair and cloned again. The PCR product cloning vector pGEM-T Easy (Promega Corporation, Madison, Wis.,

Table 1 Degenerate primers used to amplify putative cotton resistance gene analogues (RGAs) from the root-knot nematode resistance source cultivar Auburn 634 RNR. Gene products used to design the degenerate primers and the motifs of the gene products are listed along with the pathogen resistance conferred by the resistance (R) gene (Chen et al. 1998)

Primer	Sequence (5′–3′)	Resistance gene	Reference
Pto kin 1	GCATTGGAACAAGGTGAA	Tomato *Pto* gene (protein kinase)	Chen et al. 1998
Pto kin 2	AGGGGGACCACCACGTAG		
NBS-F1	GGAATGGGNGGNGTNGGNAARAC	*Arabidopsis thaliana RPS2* gene (nucleotide-binding site) and tobacco *N* gene (nucleotide-binding site)	Yu et al. 1996
NBS-R1	YCTAGTTGTRAYDATDAYYYTRC		
RLK for	GAYGTNAARCCIGARAA	Wheat *Lr10* gene (serine/threonine kinase subdomains II to VIII)	Feuillet et al. 1997
RLK rev	TCYGGYGCRATRTANCCNGGITGICC		
S2	GGIGGIGTIGGIAAIACIAC	*A.thaliana RPS2* gene (leucine-rich repeat region) and tobacco *N* gene (leucine-rich repeat region)	Leister et al. 1996
AS3	IAGIGCIAGIGGIAGICC		
CLRR for	TTTTCGTGTTCAACGACG	Tomato *Cf9* gene (leucine-rich repeat region)	Chen et al. 1998
CLRR rev	TAACGTCTATCGACTTCT		
RLRR for	CGCAACCACTAGAGTAAC	*A.thaliana RPS2* gene (leucine-rich repeat region)	Chen et al. 1998
RLRR rev	ACACTGGTCCATGAGGTT		
NLRR for	TAGGGCCTCTTGCATCGT	Tobacco *N* gene (leucine-rich repeat region)	Chen et al. 1998
NLRR rev	TATAAAAAGTGCCGGACT		
XLRR for	CCGTTGGACAGGAAGGAG	Rice *Xa21* gene (leucine-rich repeat region)	Chen et al. 1998
XLRR rev	CCCATAGACCGGACTGTT		

USA) was used to clone the RGA amplification products. Competent *Escherichia coli* DH5α cells were generated and transformed according to Hanahan (1983).

Sequencing and sequence analysis

For each *E. coli* transformation, five to ten white colonies were picked for sequencing. RGA clones were sequenced using a Beckman Coulter CEQ 8000 Genetic Analysis System (Beckman Coulter, Fullerton, Calif., USA). All clones were sequenced in both directions, using SP6 and T7 primers and the CEQ DTCS Quick Start Kit (Beckman Coutler). Sequence editing was performed using the program ChromasPro (http://www.technelysium.com.au/ChromasPro.html).

Multiple alignments and tree construction

RGA nucleotide sequences were aligned using the program Clustal X (Thompson et al. 1997), with default cost settings for opening and extending gaps. The neighbor-joining method (Saitou and Nei 1987) was used to generate a phylogenetic tree of the RGA nucleotide sequences. RGA sequences were initially divided into groups and sub-groups according to genetic distance. The RGA sequences in the sub-groups of each group were then analyzed by multiple sequence alignment using Clustal X. Sequences within each sub-group were aligned using the Clustal X program, and the multiple sequence alignment editor program GeneDoc (Nicholas et al. 1997) was used to generate a consensus sequence from the sub-group alignments. In some cases, sub-groups were unable to generate a suitable consensus sequence, due either to nucleotide sequence length constraints or mismatches and/or gaps in the alignment. In the latter case, PCR primers were designed for each individual nucleotide sequence within the sub-group. Following primer design, genomic DNA isolated from cotton cultivars Auburn 56 and Auburn 634 RNR were used as PCR templates to determine functionality of the RGA-STS primers. The primer design program FastPCR (Kalendar 2004) was used to design RGA-STS primers from the consensus sequence of each sub-group (Table 2).

Predicted amino acid sequences of cotton R proteins within RGA sub-groups that were mapped to homeologous chromosomes 12 and 26 of cultivated tetraploid cotton were aligned to compare conserved motifs present in the predicted cotton R proteins. In many cases, the deduced amino acid sequences of the cotton RGAs identified in other studies were already provided in GenBank. This information allowed for prediction of the appropriate reading frame for cotton RGAs isolated in this study. Once the cotton RGA coding sequences were deduced, the genes were translated in the appropriate reading frame and aligned using the program Clustal X (Thompson et al. 1997).

Linkage map construction

To maximize number of polymorphic RGA-STS markers and map them with existing DNA markers that are

Table 2 List of primers generated R-gene multiple sequence alignments. Cotton RGA sequences within each sub-group were aligned using Clustal X and the sub-group alignment consensus sequence used to generate forward and reverse primers. Sub-groups that contained only one sequence for primer design are indicated by an *asterisk*. In the case that a sub-group alignment did not produce a consensus sequence that could be used to generate primers, all sequences within the sub-group were used to design primers. Multiple primers from the same sub-group are indicated by the same sub-group name followed by a *dash* and *numerical designation*. Clone IDs and accession numbers are listed for sub-group sequences with sequence-specific primers

RGA sub-group	Forward (5′–3′)	Reverse (5′–3′)	Product (bp)	Temp (°C)	Clone ID and GenBank accession no.
1A	gctggaaggcgtgattgg	tggcacagtctcgcaacc	172	53.4	
1B	tgttcatgaagggaatgc	gattcgataagatcggagc	184	50.9	
1C	ttgcttgacggatgcttc	attgaaaagccgaagtgc	270	51	
1D*	tcttgctgacgttcgagg	aatcatctttcggcgttg	362	51.5	NMGhRGA-4
1E	tcttgctgacgttcgagg	cctaaaccaaaccaatcacg	209	50.5	
1F*	ttcacgaagggaatgccata	ctccctaaaccaaaccaatc	130	52.4	NMGhRGA-9H
1G*	tgtcacctcattttgaagga	tgaaatcatcttttggcattg	412	55.9	NMGhRGA-12
1H	tccttgctgacgttcgag	taaggggaggccaccagc	423	52.5	
1I	ctttctggccgatgttcg	caaaccaagctcgcctgc	228	51.1	
1J	gaagacgactcttgcgag	tacaccgaaccaatcacg	293	51.8	
2A*	aacttccatcgtgcttgg	tcatcgggcactgaggac	168	52.6	GhFungusRes-AF084206
2B*	gggcctcttgcatcgtgg	tcctaacttggccgaaac	136	50	NMGhRGA-16F
2C	agggcctcttgcatcgtg	tcctagcttggccgaaac	137	51.6	
2D*	tgcatcgtggtggaaaagaa	aaaagtgccggactcaatga	147	52	NMGhRGA-16E
3A	ccatcatgaagcacatcc	ggtagacctgcacattcc	473	52.9	
3B	cttcgatattcgaaggctac	tttgttgttggctcaagg	192	50.8	
3C	ttgcagaggcaatgaacg	ccagcacaacgctcaacg	350	52.6	
3D*	ggaaaatggcagtgcgac	agcaaagggcaggatacc	170	52.4	GhRGA_C9NBS-AF469090
3E	atttgggcagttgcatcg	ccaggggaagtcctgcac	438	54.8	
3F	acaagttcagcaccacac	tccaccgcacctttcagc	459	53.6	
3G	ctgtggttatggctgttg	gagggccagtggaagacc	432	53.4	
3H*	actggtcctgagcctgac	acatagtcgccaacagtg	634	53	GhFusariumInf-CD486128
3I*	attttcgtgttcaacgacgg	ttgggaagtctcctgtgattt	656	52	NMGhRGA-11J
4A*	tgaggttggctcagaagc	acacttggtggcaatctg	321	54	GbNBS149–3-AY244700
4B*	cattgtgaccactcgaag	gattctcattaggtttgg	158	47.5	GhRGA_C9NBS-AF469074
4C*	taagggtctgctgattgg	atctggatgcatgcttgg	212	52.6	GhRGA_C9NBS-AF469079
4D	ccattatgaggacagttcc	tgaccattgcttcactgg	142	50.1	
4E	tcaacaatgggtgctttc	tcctccaaagtcccttgc	105	50.2	
5A	attcccccacagcaagtc	tgttgacaacacgttcag	164	50.9	
5B	tcttcgacaacgagcctc	atccctgagccgagtgac	204	53.6	
5C*	agcgaggggtagacctgc	ccatcatgaagcacatcc	480	53	GbNBS12–4-AY244675
5D	tccagttccaggggctgc	atgttcagcaccccacag	348	54.2	
5E	caaaccagcacaacgctc	ttgcagaggcaatgaacg	354	52.5	
5F	tccagttccaggggctgc	gttcagcaccccacagag	346	54.5	
6A-1*	acatgcccaagcttagggcac	cacctcacaaatgctcgatgg	312	58.6	GhFusariumInf-CD486153
6A-2*	tcgaacctgtgggtgatgctc	cgcggggatctttctcgaacg	258	60.8	GhW11WiltRescDNA-CF075615
6B*	ttgcatcgtatggttctcgt	taaaaagtgccggactgaag	321	52	NMGhRGA-14D
6C-1*	tcactggtcgagttccagcag	tgtgcaccaatggggtttacg	401	60	GhLRR-AY040533
7A-1*	gaccattgcttcactggttgc	gcatccattatgaggacagttcc	145	56.9	GhRGA_C9NBS-AF469086
7A-2*	tgctttaccgattgccatcag	tctgctgattgggatgagctc	226	58.8	GhRGA_C9NBS-AF469080
7A-3*	actccctaaagcatgttgtgc	actggatggctcttcgcagtc	174	57.3	GhRGA_C9NBS-AF469081
7A-4*	caaagtcccttgctcccaatg	caatgggtgctttccatgctc	97	56.2	GhRGA_C9NBS-AF469075
7B*	aaattgggaagtctcctgca	ttcgtgttcaacgacgtactg	413	53	NMGhRGA-12I
8A-1*	taatcggcacgaatcagc	actgcagaacttcggacc	477	54	GhFusariumInf-CD486375
8A-2*	ttgaagaagcctttgcggaac	gccatcccttaatgtttgcgac	100	55.6	GhRGA_C9NBS-AF469073
8B*	tagtggaagcagtagacc	aaatacaaagctccctcc	397	52.9	GhW7WiltRescDNA-CF075622
8C*	gaggtctgccctggggattcc	tgaggtttccccaaggcttcc	255	60.7	GhW3WiltRescDNA-CF075619
8D*	gggtggatggtgttcctg	cctatggtcgacctgcag	818	55.1	GhW6WiltRescDNA-CF075621
8F*	atcagagctgcgcccgag	tgacaagcgtgacggcac	384	54.6	FiberNBS/LRR-AW561917
8G-1*	tgtagctaaccggtatcc	tgatcctagggagtcctc	143	51.6	FiberNBS/LRR-AW587467
8G-2*	gtggggaaaacatggctagt	tgattttgcgatccgattgc	237	53.2	NMGhRGA-4G
9B*	tgctccctaaaccgaacc	tgggaagacgactcttgc	303	52.3	
9C*	aatcatctttcggcgttg	tggggaagacgacccttg	451	51.9	NMGhRGA-10D
9D	tctaagtgcatcgtctgg	aggcaaaagctttcttgc	347	51.8	
9E	tttgaaggcgaggggggag	ccataattagccacaggttg	324	51.7	
9F	tgagggcgaaggggagtc	ccataattagccacaggttg	322	52	
9G*	gggagaccaccagcataattt	aagacgactcttgcaagggtt	491	54	NMGhRGA-11
9H-1*	cagtgaaactgttggagc	tatggctgtggtttggtc	381	51.8	GhW15WiltRescDNA-CF075617
9H-2*	tgaatgtagctaaccggtatcctg	ctgatgatcctagggagtcctc	151	58.1	NMGhRGA-5D
10–1*	ttgcagtgcagtgcggaatcc	gtcccaccatggttgagcgag	200	58	GhW13WiltRescDNA-CF075616
11*	tccagttgggaatggctccag	tccggtgactatcacggatgc	157	60.9	*Gossypium hirsutum* SINAH1 AF175124

on known chromosomes, genomic DNA isolated from two closely related cultivated cotton, Pima (*G. barbadense*) 3–79, and the Upland cotton (*G. hirsutum*) line NM24016 (Cantrell and Davis 2000) was used. An RIL population from a cross between the two lines was previously developed to construct a linkage map in this laboratory (Lu et al. 2004). RGA-STS primers that generated polymorphic bands were used to genotype the RIL mapping population that consisted of 72 lines, of which 40 lines were randomly selected for mapping. The program Mapmaker Macintosh, version 2.0 (Lander et al. 1987), was used to generate linkage maps of the DNA markers that were polymorphic between the parents 3–79 and NM24016, using a likelihood odds threshold of 3.0 and a recombination fraction of 0.4. PCR reactions for the RGA-STS primers were performed in 20 µl with the following concentrations: 0.5 µM primers, 1 ng/µl genomic DNA template, 1.5 mM $MgCl_2$, 0.4 mM dNTPs, and 0.05 U/µl AmpliTaq DNA polymerase (Applied Biosystems). Thermal cycling conditions were 5 min at 94°C for initial denaturation followed by 30 cycles at 94°C for 1 min, 50–58°C for 1 min (primer dependent, see Table 2), and 72°C for 1 min. A final extension was performed at 72°C for 5 min. Annealing temperatures for RGA primer pairs are listed in Table 2. All PCR reactions were performed in a PE Applied Biosystems GeneAmp PCR System 9700 (Applied Biosystems). In some cases where no amplification product was detected for a RGA primer pair, PCR conditions were optimized using a PerkinElmer Cetus DNA Thermal Cycler 480 (Perkin Elmer, Boston, Mass., USA).

Simple sequence repeats (SSR) and sequence related amplified polymorphisms (SRAP) markers for the RIL population were as reported in Lu et al. (2004).

Screening for RGA-STS polymorphisms in RKN-resistant and -susceptible NILs

Genomic DNA from the NILs Auburn 56 and Auburn 634 RNR, Stoneville 213 and M-249 RNR (Shepherd et al. 1989, 1996), Deltapine 61 and M-240 RNR (Shepherd et al. 1989, 1996), and Coker 201 and M-120 RNR (Shepherd et al. 1989, 1996) was isolated as previously described. Genomic DNA was used as templates in PCR reactions with RGA-STS primers (Table 2) to identify RGA polymorphic markers specific for the RKN-resistant variety of each NIL pair. The PCR reaction conditions were the same as those previously described for the RGA-STS linkage map construction.

Results

Amplification of RGAs from Upland cotton

The eight degenerate primer pairs (Table 1) amplified multiple bands from both Auburn 56 and Auburn 634 RNR. Major fragments were present in the two NILs, indicating a robust and consistent PCR amplification, using the primers and PCR conditions as described previously; however, in some cases, polymorphic but minor bands were observed in only Auburn 634 RNR (Fig. 1). Therefore, the bands from only Auburn 634 RNR were excised and purified from the agarose gel for cloning. Of a total of 165 clones isolated, sequence analysis revealed 57 to be novel nucleotide sequences, many of which share a high degree of homology to RGA nucleotide sequences previously isolated from *G. hirsutum* and *G. barbadense*. A total of 36 RGA nucleotide sequences were submitted to GenBank. Of the remaining 21 nucleotide sequences, two were identical to a *G. hirsutum* SINAH1 homologue sequence (Jaradat 1999) already deposited in GenBank. The remainders were deemed too short (52–89 bp in length) with not enough homology information available to infer functionalities or coding sequences. These sequences were included in the RGA nucleotide sequence cluster analysis, but were not assigned GenBank accession numbers (Fig. 2).

Cluster analysis of RGA nucleotide sequence

A cluster analysis was performed with 57 RGA nucleotide sequences isolated in this study, in addition to 99 cotton RGA nucleotide sequences present in GenBank. The cotton RGA nucleotide sequences clustered into 11 groups and 56 sub-groups. Of the 11 groups, RGAs identified in the present study formed five unique groups (2, 3, 6, 8, and 11), while groups four, five, and ten are exclusively composed of RGAs deposited previously in GenBank. Group 11 of the cluster analysis contains the two nucleotide sequences identical to the *G. hirsutum*

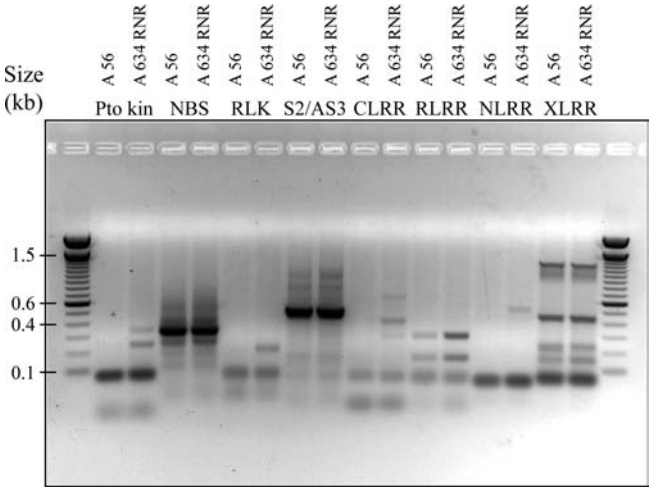


Fig. 1 Putative cotton resistance gene analogues (RGA) amplified using resistance (R)-gene degenerate primers were resolved on an agarose gel. RGA degenerate primers are listed *above lanes* on the gel along with the cultivar genomic DNA used as the template in the PCR reactions. Auburn 56 is represented by A 56 and Auburn 634 RNR is represented by A 634 RNR

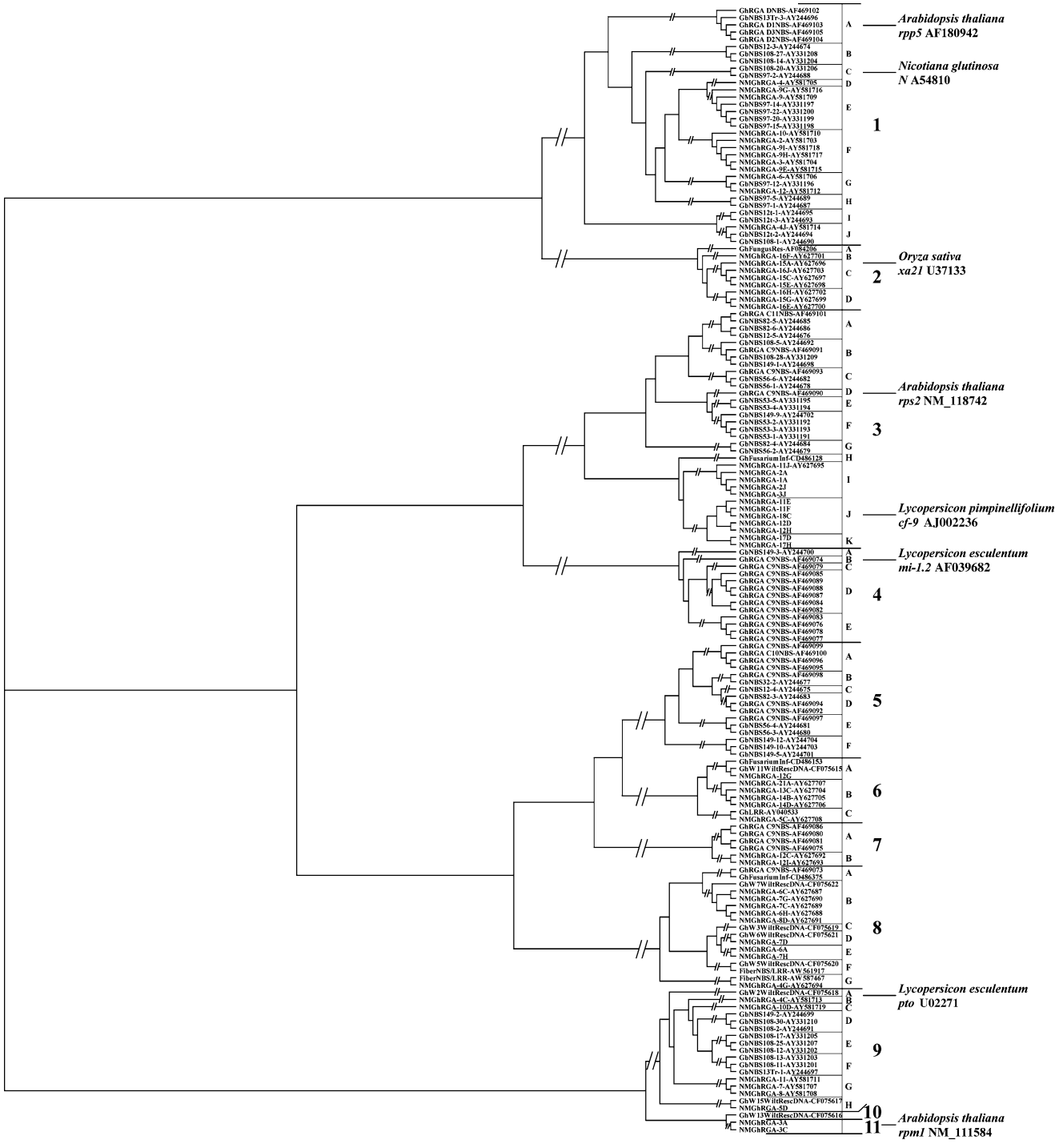


Fig. 2 Cladogram representation of a phylogenetic tree. Branch lengths are not proportional to genetic distance. RGA groups and sub-groups are shown on the *right* with sequences listed with GenBank accession numbers. RGA groups and sub-groups were determined by genetic distance and sub-group sequence homology, and are illustrated by *breaks in the branches* of the cladogram

SINAH1 homologue nucleotide sequence (Jaradat 1999). Of 56 sub-groups, 21 were from the novel RGA sequences identified in the study. The cluster analysis of the RGA sequences is represented as a cladogram in Fig. 2, with hash marks indicating groupings and sub-groupings according to genetic distance and sub-group sequence homology. It is important to note that the cladogram representation of the cluster analysis does not have branch lengths that are representative of the genetic distance among the RGA nucleotide sequences. Given the large number of RGA nucleotide sequences analyzed in this study, a cladogram is a much clearer

means of representing cotton RGA groups and sub-groups from the cluster analysis results. The RGA-STS primer pairs each amplified a prominent DNA fragment of predicted size (Table 2) as visualized on an ethidium bromide-stained agarose gel. However, in nearly all instances, other DNA fragments of varying sizes were also evident as faint, minor and prominent, major PCR products (Fig. 3).

Screening for RGA polymorphisms and mapping RGA markers

Genomic DNA from the *G. barbadense* genotype 3–79 and the *G. hirsutum* Acala-type line NM24016 were amplified for polymorphic DNA markers, using the RGA-STS primers (Fig. 4). Of the 61 RGA-STS primer pairs used in the PCR reactions, nine amplified DNA fragments that were polymorphic between 3–79 and NM24016. The polymorphic RGA markers included RGA groups 1D, 1E, 1J, 3E, 3I, 5A, 5B, 8F, and 9B (Fig. 2). The RGA-STS primers for these groups were used to amplify genomic DNA from the RIL population of 3–79 × NM24016. Interestingly, all nine RGA-STS markers were mapped to homeologous chromosomes 12 and 26 (Fig. 6). RGA groups 1E, 1J, 3I, 5A, 5B, and 9B were mapped to a 25.6-cM region that includes the SSR marker BNL 1679, which is located on the chromosome 12 of the A sub-genome (Liu et al. 2000). Six SSR were also mapped on chromosome 12. RGA groups 1D, 3E, and 8F were mapped to a 40.5-cM region that includes the SSR marker BNL 341, which is located on the chromosome 26 of the D sub-genome (Lacape et al. 2003). In addition to the nine RGA-STS markers mapped to homeologous chromosomes 12 and 26 (Fig. 6), the linkage map of chromosome 12 includes seven SSR

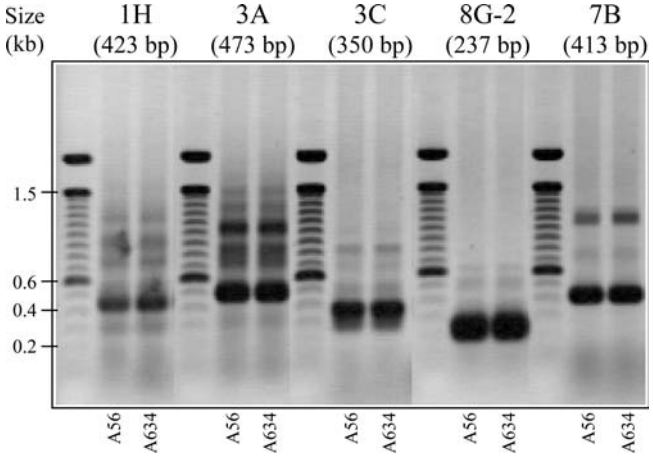


Fig. 3 Genomic DNA from cotton cultivars Auburn 56 and Auburn 634 RNR amplified with RGA-sequence-tagged sites (STS) primers and resolved on an agarose gel. The figure is a composite of multiple gels and illustrates the multiple bands amplified using the RGA-STS primers. RGA sub-groups (Fig. 2) from which the primers were designed are listed *above the lanes*. The sizes of the predicted major amplification products for each RGA-STS primer pair are listed in parentheses. Cotton cultivars are listed at the *bottom of each lane*

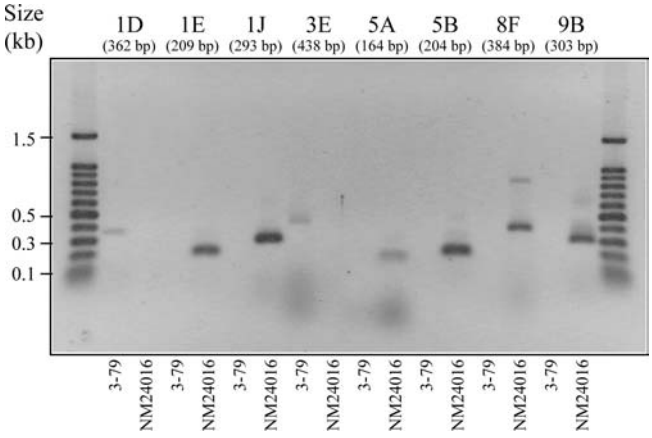


Fig. 4 Genomic DNA from cotton genotypes 3–79 and NM24016 amplified with RGA-STS primers and resolved on an agarose gel. Polymorphic bands generated by the RGA-STS primers enabled mapping of the RGA markers, using a recombinant inbred line (RIL) population from a cross between 3–79 and NM24016. RGA sub-group designations (Fig. 2) are listed at the *top of the lanes*, with predicted band sizes in *parentheses*. Cotton genotypes are listed at the *bottom of each lane*

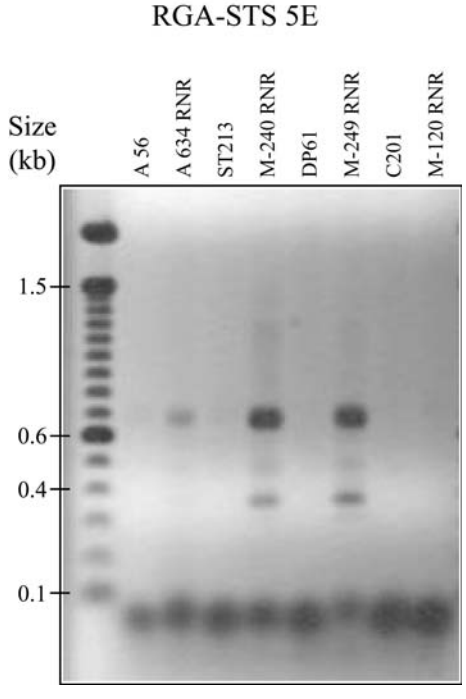
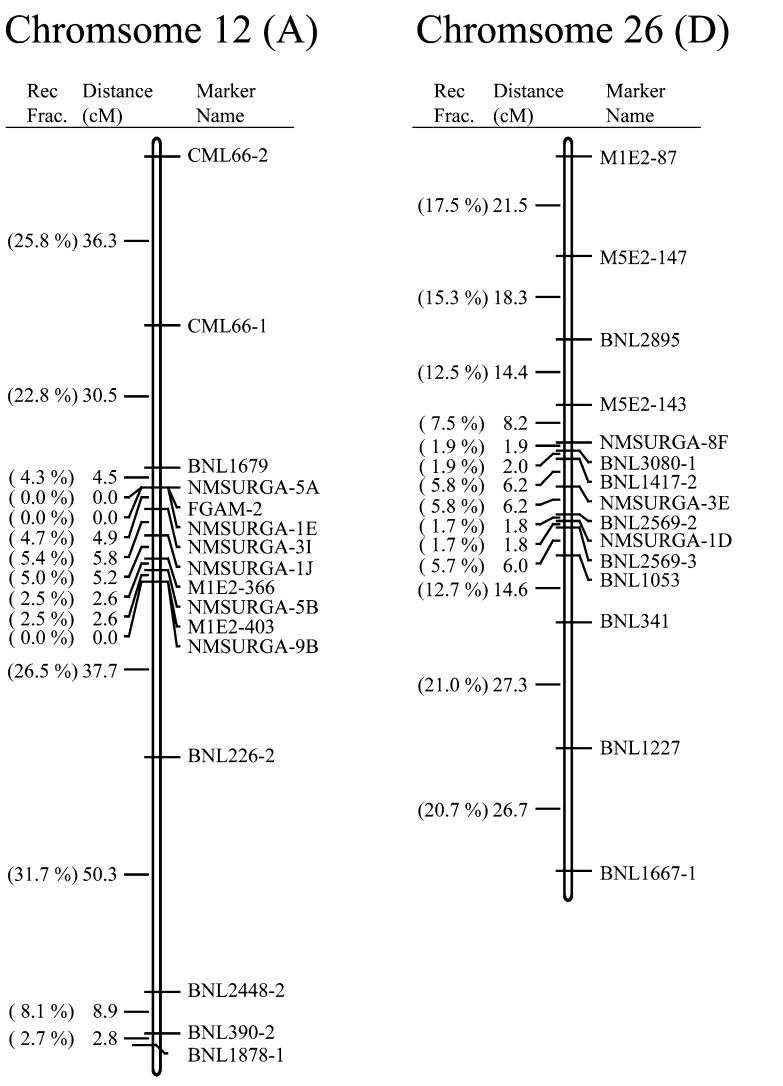


Fig. 5 Polymorphic bands amplified from genomic DNA isolated from upland cotton near-isogenic lines resistant and susceptible to root-knot nematode. The RGA-STS primer pair 5E (Table 2) amplified polymorphic bands from the resistant upland cotton cultivars Auburn 634 RNR, M-249 RNR, and M-240 RNR. No amplification products were visible from the respective susceptible recurrent parents Auburn 56, Stoneville 213, and Deltapine 61. The upland cotton cultivars from which the genomic DNA was isolated for PCR are listed at the *top of each lane*

a



b

NMSU RGA-5A	NMSU RGA-1E	NMSU RGA-1J	NMSU RGA-3I	NMSU RGA-5B	NMSU RGA-9B	NMSU RGA-8F	NMSU RGA-3E	NMSU RGA-1D
AF469095	AY331197	AY244690	AY627695	AF244677	AY581713	AW561917	AY331194	AY581705
AF469096	AY331198	AY244694		AF469098		CF075620	AY331195	
AF469099	AY331199	AY581714						
AF469100	AY331200							
	AY581709							
	AY581716							

Fig. 6 **a** Cotton R-gene linkage maps of nine RGA-STS markers mapped to homeologous chromosomes 12 and 26 of an RIL of tetraploid cotton. The program Mapmaker was used to generate the linkage maps with a likelihood odds threshold of 3.0 and a recombination fraction of 0.4. Distance between markers is listed in centiMorgans, with the recombination fraction in *parentheses*. The simple sequence repeats (SSR) marker BNL1679 is located on chromosome 12. The SSR markers BNL341 and BNL1227, are located on chromosome 26. **b** GenBank accession numbers of RGA sequences from sub-groups (Fig. 2) used to generate RGA-STS primers (Table 2). The sequence accession numbers are listed *below the corresponding RGA-STS marker* from the linkage groups

markers, two SRAP markers (Li and Quiros 2001; Lu et al. 2004), and an STS marker for a putative cotton phosphoribosylformylglycinamidine synthase gene identified in this laboratory by representational difference analysis (Baldocchi and Flaherty 1997). The linkage map of chromosome 26 includes nine SSR markers and three SRAP markers (Lu et al. 2004).

To identify potential polymorphic RGA-STS markers associated with RKN resistance, genomic DNA from the four pairs of cultivated cotton NILs that are resistant or susceptible to RKNs were screened with the RGA-STS primers (Table 2). Two RGA-STS primers amplified polymorphic bands between resistant and susceptible NILs (Fig. 5). The putative DNA markers

(RGA-STS 5E and RGA-STS 6C-1) were present as RGA-STS primer amplification products only in PCR reactions using genomic DNA templates isolated from the RKN-resistant lines Auburn 634 RNR, M-240 RNR, and M-249. No RGA-STS primer amplification products were evident from PCR reactions using genomic DNA templates isolated from the respective RKN-susceptible cultivars Auburn 56, Deltapine 61, and Stoneville 213. However, RGA-STS primers designed from the clones isolated from the previously minor polymorphic bands (Fig. 1) did not yield any polymorphism among the NILs.

Sequence conservation of cotton RGAs mapped to homeologous chromosomes 12 and 26 of tetraploid cotton

To further examine the sequence relationship among the mapped RGAs on homeologous chromosomes 12 and 26, a total of 20 R-protein sequences representing eight RGA-STS markers were used for a multiple alignment. The alignment of RGA-predicted amino acid sequences revealed a high degree of conservation in two R-protein functional sequence motifs (Fig. 7). Predicted protein sequences in RGA sub-groups RGA-1E and RGA-1J

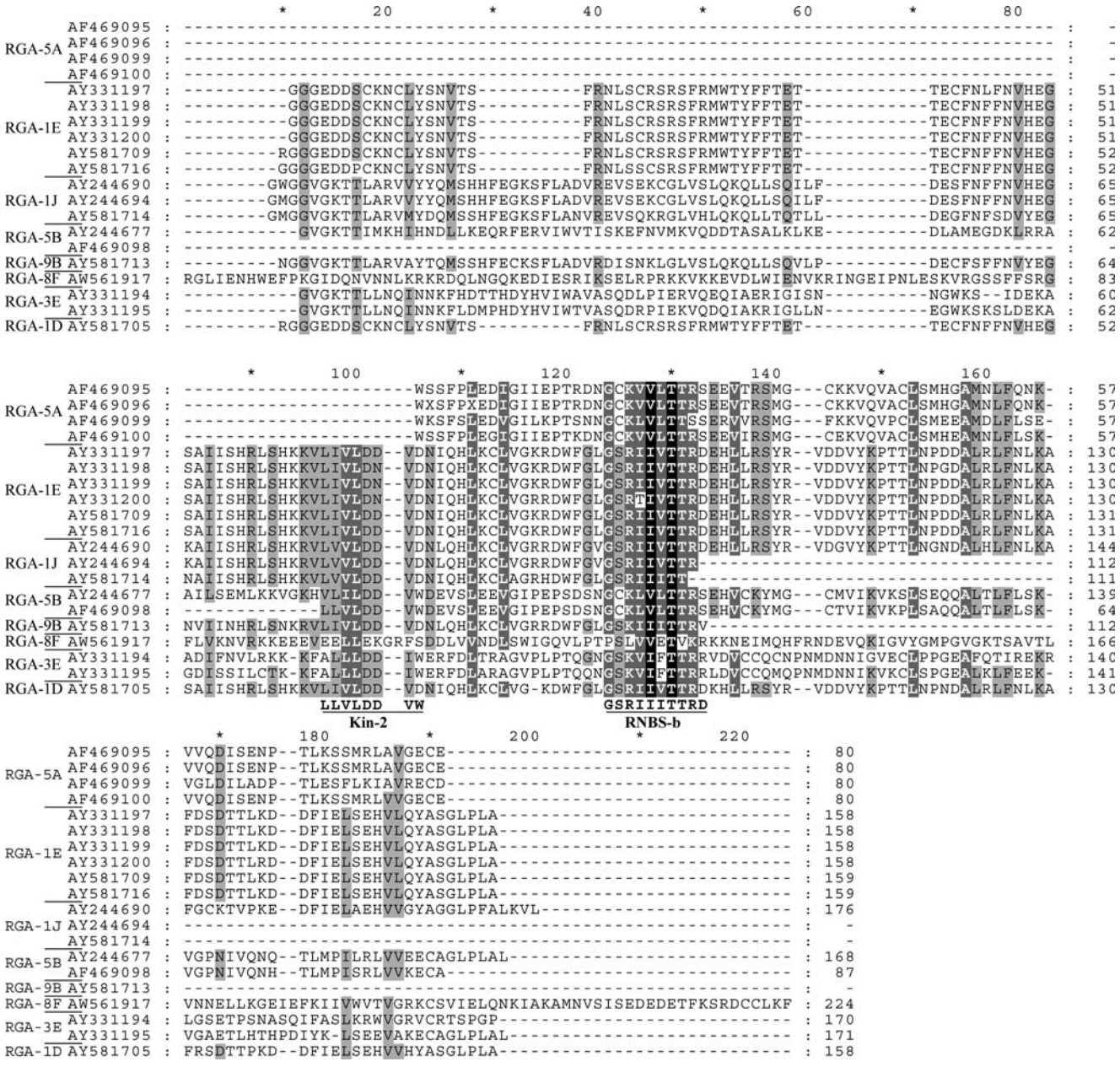


Fig. 7 Multiple alignment of predicted R-gene amino acid sequences from RGA-STS markers shown on the linkage maps in Fig. 6. Conserved R-protein motifs Kin-2 and RNBS-b are indicated *below the multiple alignments*. Mapped RGA-STS marker designations (Fig. 6) of aligned sequences are listed to the *left of the alignment*

contains homeologous sequences from *G. barbadense* and *G. hirsutum*, indicating that R-protein sequences in the multiple alignments are possibly paralogues and/or orthologues. Deduced cotton RGA proteins from all the mapped sub-groups except RGA-5A contain the protein motifs Kin-2 and RNBS-b, as described in Meyers et al. (1999). RGA sub-group RGA-5A contains the conserved RNBS-b motif, but is lacking the Kin-2 motif. However, all amino acid sequences present in the multiple alignments are incomplete at the N terminus. In the case of sub-group RGA-5A, the missing N-terminal sequence data is located C terminal to the location of the Kin-2 motif of the other R-protein sequences in the multiple alignments, so the Kin-2 motif could theoretically exist in the RGA-5A sub-group sequences. Amino acid sequence and nucleotide sequence conservation among the mapped cotton RGAs, particularly within each RGA sub-group, corresponds with observations in other species that R genes are clustered in the genomes of plants (Michelmore and Meyers 1998; Leister 2004).

Discussion

Resistance genes of plants, both dicotyledonous and monocotyledonous, comprise a significant portion of the genome, represented by approximately 200 genes in the entire genome of *Arabidopsis thaliana* (Meyers et al. 2003). Recent analysis of the rice genome identified approximately 500 non-TIR/NBS/LRR genes that account for 1% or more of the entire rice genome (Monosi et al. 2004; Zhou et al. 2004). Plant R genes are classified based on amino acid motifs present in the R-gene products that are presumed functional in pathogen recognition and elicitation of a defense response. Many of these sequences differ only slightly by substitutions, deletions, and/or insertions at the nucleotide level. Theories on R-gene duplication in the plant genome include tandem and ectopic gene duplication events and recombination (Meyers et al. 2003; reviewed in Leister 2004) This R-gene replication represents the co-evolutionary ''arms race'' between plants and the multitude of pathogens they must face, and is considered analogous to vertebrate T-cell receptor, major histocompatibility complex, and immunoglobulin genes (reviewed in Michelmore and Meyers 1998). In most plant species studied to date, R genes are known to cluster to regions on chromosomes that represent loci involved in specific pathogen recognition, and ultimately, resistance to that particular pathogen (Michelmore and Meyers 1998). In many cases, this resistance phenotype manifests itself as the well-known hypersensitive response, followed by systemic acquired resistance (Durrant and Dong 2004).

In this study, one of our major concerns was the development and mapping of RGA-STS markers derived from the RGA-STS primer pairs (Table 2) designed from the cluster analysis (Fig. 2). In addition to a clearer representation of the RGA sub-groups assigned to the 156 tetraploid cotton RGAs, a cladogram was used to represent the grouping and sub-grouping, since the purpose of this study was not to infer phylogenetic relationships among the cotton RGAs. Groups were assigned by genetic distance, but sub-groups were determined by genetic distance and the ability to generate sub-group RGA consensus sequences from which primers could be designed. Therefore, further confirmation of the neighbor-joining tree created by the Clustal X multiple alignment by bootstrapping (Felsenstein 1985) the tree was unnecessary.

The cotton RGA-STS primer pairs (Table 2) designed from the consensus sequences or individual sequences of RGA sub-groups (Fig. 2) were developed to amplify a major DNA fragment of specific size. However, especially in the case of consensus sequence derived RGA-STS primers, multiple bands are evident as major and minor amplification products for each pair of RGA-STS primers (Fig. 3). In all cases, the major amplification products of predicted sizes were present when PCR products were resolved on an agarose gel. Since the RGA-STS primers are highly specific, we presume that the other bands are also cotton RGA sequences representing an R-gene family containing the NBS/LRR R-protein motifs. This presumption conforms to observations that many R-gene families cluster in plant genomes to loci associated with resistance to a specific pathogen or pathogens (Graham et al. 2002; Huettel et al. 2002; reviewed in Michelmore and Meyers 1998; Leister 2004). Unidentified cotton RGAs that contain insertions and/or deletions at the nucleotide level, yet still retain homology in the gene sequences from which the RGA-STS primers were designed, would amplify PCR products of varying sizes from the same genomic DNA template. The high degree of sequence homology among the protein sequences mapped to tetraploid cotton homeologous chromosomes 12 and 26 (Fig. 6) is also indicative of R-gene clustering; however, any pathogen resistance conferred by these loci are presently unknown. Gaps in the multiple alignment of protein sequences from different RGA sub-groups (Fig. 7) may result from unequal crossover events that give rise to new copies of R genes in the cluster. Variations in the size of genes duplicated in this manner could also lead to multiple PCR products of varying sizes as demonstrated in Fig. 3.

Previous studies on RGAs isolated from tetraploid cotton have focused on a limited number of PCR amplification products for isolation and sequencing of putative R genes. Tan et al. (2003) utilized R-gene degenerate primers designed from the NBS motifs of the tobacco N protein, *Arabidopsis* RPS2 protein, and the flax L6 protein to amplify and clone PCR products in the 250-bp size range. The 250-bp amplification products were representative of the genomic DNA sequences transcribed and translated into the region between the kinase-2 and GLPL motifs of plant R proteins. This approach enabled cloning of 33 putative cotton RGAs containing the highly conserved NBS R-protein motif

(Tan et al. 2003). A more recent study by He et al. (2004) utilized degenerate primers designed from NBS and membrane-spanning motifs representing TIR/NBS/LRR and non-TIR/NBS/LRR classes of R proteins. Cloning of a 560-bp fragment allowed for identification of 61 unique sequences containing high similarity to R genes already deposited in GenBank. In addition to cloning and identification of novel cotton RGAs that have been deposited into the GenBank public sequence repository, we were also interested in utilizing the cultivated tetraploid cotton RGAs already reported therein. Amplification of putative R-gene sequences, using degenerate primers designed from R-protein motifs representing a variety of characterized plant R proteins (Table 1), enabled amplification of a large number of novel putative cotton RGAs (Fig. 2). Combining the accumulated GenBank RGA sequence data from *G. hirsutum* and *G. barbadense* provided a detailed view of RGA nucleotide and protein sequence homology by multiple alignment and/or cluster analysis. The RGASTS primers designed from this study (Table 2) represent possible cultivated cotton R-gene markers that may be associated with potentially a vast number of resistance phenotypes.

The proven method of utilizing degenerate primers designed from known plant R-gene products was used in this study to isolate and identify novel R genes from the cultivated cotton species *G. hirsutum* cultivar Auburn 634 RNR. Auburn 634 RNR is the source of RKN resistance in tetraploid cotton (Shepard et al. 1989); however, no gene(s) directly involved in the resistance phenotype has/have been identified. Likewise, tentative DNA markers linked to the resistance phenotype are mapped too distantly to RKN resistance to justify utilization in traditional breeding programs (Bezawada et al. 2003). One of the currently available commercial cotton cultivars that processes resistance to RKNs is the Acala cultivar NemX (Oakley 1995). The Acala NemX shows significant resistance to RKNs, and rotations using NemX can reduce nematode field populations for subsequent growing seasons with susceptible cultivars (Ogallo et al. 1997). Utilization of NemX is limited to the southwestern United States, mainly in California's San Joaquin Valley. Most of the commercial cotton varieties grown in other regions of the United States remain susceptible to RKNs. As previously mentioned, the monetary loss due to RKNs and faced yearly by cotton farmers is quite significant.

Our previous studies using two segregating populations derived from crosses between Sure-Grow 747 and Auburn 634 RNR, and Sure-Grow 747 and M-240 RNR, indicate that one or two genes are responsible for RKN resistance (Zhang et al. 2004). Segregation of the resistance phenotype in both mapping populations fit a 3:1 ratio by chi-square analysis, indicating one dominant RKN resistance gene. In the F_2 mapping population derived from a cross between Sure-Grow 747 and M-240 RNR, segregation of the resistance phenotype also fit a 13:3 ratio by chi-square analysis, indicating possible epistatic interaction between a dominant and a recessive gene conferring RKN resistance (Zhang et al. 2004).

Our interest is in developing markers associated with RKN resistance for MAS and identification of gene(s) directly involved in RKN resistance in cultivated tetraploid cotton. Annual crop losses due to nematodes are estimated at approximately $8 billion in the United States and approximately $78 billion worldwide (Barker et al. 1998). The two main species of nematodes that are associated with cotton farming revenue losses in the United States are the RKN (*M. incognita*) and the reniform nematode (*Rotylenchulus reniformis*). However, distribution of these two species of nematode is not uniform in the continental United States. RKNs are a problem predominantly in the southwestern United States, while reniform nematodes are currently found in the southeastern United States and spreading westward. In addition to yield reductions caused directly by RKN damage to the plant root system, infestation also causes increased susceptibility to fungal pathogens such as Fusarium wilt. Identification of DNA markers and/or gene(s) involved in resistance to RKNs is of great interest in the cotton farming industry. To date, no DNA markers have been identified in cultivated cotton that are linked close enough to the RKN resistance phenotypes to facilitate an MAS breeding program. The putative RKN resistance-associated DNA markers identified in this study (Fig. 5), using two of the RGA-STS primer pairs, may prove useful in MAS once the association of these markers with RKN resistance is determined. The RGA-STS primers developed in this study (Table 2) may prove useful tools for researchers, not only in the field of RKN resistance, but in the development of cotton DNA markers and identification of genes associated with resistance to a variety of pathogenic organisms.

Acknowledgements The research presented here was funded in part by Cotton Incorporated, Cary, N.C., USA; the United States Department of Agriculture through the Southwest Consortium on Plant Genetics and Water Resources; and the New Mexico Agriculture Experiment Station.

References

Baldocchi RA, Flaherty L (1997) Isolation of genomic fragments from polymorphic regions by representational difference analysis. Methods 13:337–346

Barker KR, Pederson GA, Windham GL (eds) (1998) Plant nematode interactions Madison, American Society of Agronomy, Crop Science of America, Soil Science Society of America

Bezawada C, Saha S, Jenkins J, Creech R, and McCarty J (2003) SSR marker(s) associated with root-knot nematode resistance gene(s) in cotton. J Cotton Sci 7:179–184

Blasingame D, Patel MV (2004) Cotton disease loss estimate committee report. In: Proceedings of the 2004 beltwide cotton conferences. National Cotton Council of America, Memphis, pp 459–460

Borrone JW, Kuhn DN, Schnell RJ (2004) Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. Theor Appl Genet 109:495–507

Cantrell RG, Davis DD (2000) Registration of NM24016, an interspecific-derived cotton genetic stock. Crop Sci 40:1208

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

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

Durrant WE, Dong X (2004) Systemic acquired resistance. Annu Rev Phytopathol 42:185–209

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Feuillet C, Schachermayr G, Keller B (1997) Molecular cloning of a new receptor-like kinase gene encoded at the Lr10 disease resistance locus of wheat. Plant J 11:45–52

Graham MA, Marek LF, Shoemaker RC (2002) Organization, expression and evolution of a disease resistance gene cluster in soybean. Genetics 162:1961–1977

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580

He L, Du C, Covaleda L, Xu Z, Robinson AF, Yu JZ, Kohel RJ, Zhang H-B (2004) Cloning, characterization, and evolution of the NBS-LRR-encoding resistance gene analogue family in polyploid cotton (*Gossypium hirsutum* L.). Mol Plant–Microbe Interact 17:1234–1241

Huettel B, Santra D, Muehlbauer J, Kahl G (2002) Resistance gene analogues of chickpea (*Cicer arietinum* L.): isolation, genetic mapping and association with a Fusarium resistance gene cluster. Theor Appl Genet 105:479–490

Jaradat TT (1999) Studies of cotton cDNAs encoding a seven in absentia homolog, a potential DNA binding and a cell wall protein, PhD Dissertation, Texas Tech University, Lubbock

Kalendar R (2004) FastPCR, PCR primer design, DNA and protein tools, repeats and own database searches program. Available via http://www.biocenter.helsinki.fi/bi/bare-1_html/fastpcr.htm

Kuhn DN, Heath M, Wisser RJ, Meerow A, Brown JS, Lopes U, Schnell RJ (2003) Resistance gene homologues in *Theobroma cacao* as useful genetic markers. Theor Appl Genet 107:191–202

Lacape JM, Nguyen TB, Thibivilliers S, Bojinov B, Courtois B, Cantrell RG, Burr B, Hau B (2003) A combined RFLP-SSR-AFLP map of tetraploid cotton based on a *Gossypium hirsutum* × *Gossypium barbadense* backcross population. Genome 46:612–626

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174–181

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

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

Li G, Quiros CF (2001) Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in *Brassica*. Theor Appl Genet 103:455–461

Liu S, Saha S, Stelly D, Burr B, Cantrell RG (2000) Chromosomal assignment of microsatellite loci in cotton. J Hered 91:326–332

Lu Y, Zhang Jinfa, Percy RG, Cantrell RG (2004) An integrated SSR-STS-SRAP-RAPD genetic map using a recombinant inbred line population in tetraploid cotton. In: Proceedings of the 2004 beltwide cotton conferences. National Cotton Council of America, Memphis, Tennessee, pp 1156–1161

Madsen LH, Collins NC, Rakwalska M, Backes G, Sandal N, Krusell L, Jensen J, Waterman EH, Jahoor A, Ayliffe M, Pryor AJ, Langridge P, Schulze-Lefert P, Stougaard J (2003) Barley disease resistance gene analogs of the NBS-LRR class: identification and mapping. Mol Genet Genomics 269:150–161

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

Martin GB, Bogdanove AJ, Sessa G (2003) Understanding the functions of plant disease resistance proteins. Annu Rev Plant Physiol Plant Mol Biol 54:23–61

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

Monosi B, Wisser RJ, Pennill L, Hulbert SH (2004) Full-genome analysis of resistance gene homologues in rice. Theor Appl Genet 109:1434–1447

Nicholas KB, Nicholas HB Jr, Deerfield DWI (1997) GeneDoc: analysis and visualization of genetic variation. EMBNEW News 4:14

Noir S, Combers MC, Anthony F, Lashermes P (2001) Origin, diversity and evolution of NBS-type disease-resistance gene homologues in coffee trees (*Coffea* L.). Mol Genet Genomics 265:654–662

Oakley SR (1995) CPCSD Acala C-225: a new nematode-resistant Acala variety for California's San Joaquin Valley. In: Proceedings of 1995 beltwide cotton production research conference. National Cotton Council of America, Memphis, Tennessee, p 39

Ogallo JL, Goodell PB, Eckert J, Roberts PA (1997) Evaluation of NemX, a new cultivar of cotton with high resistance to *Meloidogyne incognita*. J Nematol 29:531–537

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

Shepherd RL (1974) Registration of three germplasm lines of cotton. Crop Sci 14:692

Shepherd RL, Parrott WL, McCarty J, Jenkins JN (1989) Notice of the release of nine root-knot nematode resistant germplasm lines of upland cotton, *Gossypium hirsutum* L. USDA/Mississippi Agric and Forestry Experimental Station Memo

Shepherd RL, McCarty J, Jenkins JN, Parrott WL (1996) Registration of nine cotton germplasm lines resistant to root-knot nematode. Crop Sci 36:820

Tan H, Callahan FE, Zhang X-D, Karaca M, Saha S, Jenkins JN, Creech RG, Ma D-P (2003) Identification of resistance gene analogs in cotton (*Gossypium hirsutum* L.). Euphytica 134:1–7

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

United States Department of Agriculture (2003) Cotton and wool situation and outlook yearbook. In: Evans M (ed) Market and Trade Economics Division, Economic Research Service, US Department of Agriculture

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

Zhang J, Hinchliffe DJ, Lu Y, Potenza C, Sengupta-Gopalan C, Cantrell RG (2004) Root-knot nematode resistance in auburn 634RKN: segregation and molecular mapping. In: Proceedings beltwide cotton conferences. National Cotton Council of America, Memphis, Tennessee, pp 1122–1124

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