

2000

Cloning of the Arabidopsis RTM1 gene, which controls restriction of long-distance movement of tobacco etch virus

Stephen T. Chisholm, *Washington State University*

Sunita K. Mahajan, *Washington State University*

Steven A. Whitham, *Washington State University*

Miki L. Yamamoto, *Washington State University*

James C. Carrington, *Washington State University*

Cloning of the *Arabidopsis* *RTM1* gene, which controls restriction of long-distance movement of tobacco etch virus

Stephen T. Chisholm*, Sunita K. Mahajan*, Steven A. Whitham, Miki L. Yamamoto, and James C. Carrington[†]

Institute of Biological Chemistry, Washington State University, Pullman, WA 99164-6340

Communicated by Clarence A. Ryan, Jr., Washington State University, Pullman, WA, November 16, 1999 (received for review October 1, 1999)

The locus *RTM1* is necessary for restriction of long-distance movement of tobacco etch virus in *Arabidopsis thaliana* without causing a hypersensitive response or inducing systemic acquired resistance. The *RTM1* gene was isolated by map-based cloning. The deduced gene product is similar to the α -chain of the *Artocarpus integrifolia* lectin, jacalin, and to several proteins that contain multiple repeats of a jacalin-like sequence. These proteins comprise a family with members containing modular organizations of one or more jacalin repeat units and are implicated in defense against viruses, fungi, and insects.

resistance

Plants can resist viruses by a number of mechanisms. Resistance pathways involving strain-specific recognition of a virus-encoded elicitor, through direct or indirect interaction with a corresponding resistance gene (*R* gene) product, can lead to a hypersensitive reaction involving localized cell death, activation of salicylic acid (SA)-mediated systemic acquired resistance, and limitation of the virus to initial infection foci (1). Two dominant *R* genes, *N* and *Rx*, which limit strains of tobacco mosaic virus and potato virus X, respectively, have been isolated and shown to belong to a class of proteins containing nucleotide-binding sites (NBSs) and leucine-rich repeats (LRRs) (2, 3). These NBS-LRR proteins are similar to a large number of *R* gene products that confer resistance to bacteria, fungi, and nematodes, in addition to viruses (4). Although dominant *R* gene-mediated resistance responses typically include activation of defensive genes, the molecular bases whereby viruses are suppressed by these reactions are still not clear. Equally unclear is how several other types of genetically controlled virus resistance mechanisms, including those mediated by recessive genes, operate to restrict viruses (5–7). On the other hand, mechanisms of posttranscriptional gene silencing-like responses to virus infection are becoming clear. The antiviral effect of virus-induced silencing appears to involve induction of the silencing apparatus (possibly by double-stranded RNA), maintenance of a silenced state in which RNA sequence-specific recognition and degradation results in low levels of active viral genomes, and possibly cell-to-cell and long-distance signaling of the silencing state (8–10).

The tobacco etch virus (TEV)-*Arabidopsis* model system was established to identify host genes that control susceptibility to, and defense against, viruses (11, 12). Some ecotypes (e.g., C24 and La-er) are fully susceptible, supporting replication, cell-to-cell movement, and long-distance (vasculature-dependent) movement of TEV, whereas other ecotypes (e.g., Col-0 and Ws-2) restrict TEV to inoculated leaves. Ecotypes that fail to support long-distance movement, however, support replication and cell-to-cell movement in inoculated leaves. The restriction of TEV involves neither formation of hypersensitive reaction lesions nor activation of systemic acquired resistance. This natural variation in susceptibility is conditioned by the *RTM1* locus (11), where the dominant allele confers the restricted infection phenotype. A mutational analysis of Col-0 plants revealed that the

restricted infection phenotype is a multigenic trait that requires at least three loci, *RTM1*, *RTM2*, and *RTM3* (12) (unpublished data). The TEV-*Arabidopsis* system, therefore, provides an unusual opportunity to identify and functionally dissect the components of an antiviral mechanism in plants.

In this paper, isolation of the *RTM1* gene is described. This gene encodes a protein that belongs to a large “domain family,” in which a well-characterized lectin (jacalin) subunit serves as the prototype member. Proteins in this family serve diverse roles in defense against pathogens and pests.

Materials and Methods

***Arabidopsis* Inoculation and Selective Screening.** Semipurified TEV-bar virus (12) was diluted in 20 mM Tris-HCl (pH 8.0) containing 10 g/liter carborundum. Rosette leaves of 4-week-old *Arabidopsis* plants were inoculated by using an artist’s airbrush (12). At 11 days postinoculation inflorescence tissue was removed, and 0.067% glufosinate ammonium (GA) (AgroEvo, Montvale, NJ) was applied as described (12). Seven days after GA application, survivors were sprayed with 0.01% GA. Fourteen days after the initial GA application, tissue from surviving plants was collected.

Genomic DNA Extraction and Mapping. Genomic DNA was isolated from inflorescence tissue as described (11). A bacterial artificial chromosome (BAC) contig between markers ATEAT1 and F17J18-M was assembled from hybridization data generated by the *Arabidopsis thaliana* Genome Center (http://genome.bio.upenn.edu/physical-mapping/BAC_data/allhybs/allframe.html), PCR markers, DNA hybridization, and restriction pattern analysis. End sequence from BACs F2M5-T7, T12A12-Sp6, and F3F20-Sp6 (<http://genome.bio.upenn.edu/physical-mapping/bacendlist/bacends.html>) and yeast artificial chromosome yUP8H12 (GenBank accession no. AC000098) was used to generate cleaved amplified polymorphic sequence (CAPS) and dCAPS markers. Primers were designed from available sequences and used to amplify the Col-0 and Ws-2 alleles by PCR (primer sequences are available on request). The resulting PCR products were purified and sequenced. Alignments of the Col-0 and Ws-2 sequences using CLUSTAL W (13) and restriction site analysis using DNA STRIDER (14) identified polymorphisms. The CAPS and dCAPS PCR products were digested with the ap-

Abbreviations: *R* gene, resistance gene; SA, salicylic acid; NBS, nucleotide-binding site; LRR, leucine-rich repeat; TEV, tobacco etch virus; GA, glufosinate ammonium; GUS, β -glucuronidase; MBP, myosinase-binding protein; JR, jacalin repeat; BAC, bacterial artificial chromosome; CAPS, cleaved amplified polymorphic sequence; RACE, rapid amplification of cDNA ends; JA, jasmonic acid.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. AF191302).

*S.T.C. and S.K.M. contributed equally to this work.

[†]To whom reprint requests should be addressed. E-mail: carrington@wsu.edu.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

appropriate restriction enzyme (information available on request) and analyzed on agarose or polyacrylamide gels. The simple sequence length polymorphism PCR products were analyzed on 4% Metaphor agarose (FMC) gels.

Construction of Cosmid Sublibraries from BACs. High molecular weight BAC DNA was purified by cesium chloride-ethidium bromide gradient centrifugation and was partially digested with *Sau3AI*, dephosphorylated, and size-fractionated on 0.8% low melting agarose gels. Fragments of approximately 15–20 kb were eluted from gels and inserted into the *Bam*HI site of cosmid vector pSLJ75515 (15). Ligated DNA was packaged in phage heads by using commercial extracts (Stratagene Gigapack III Gold) and used to transfect DH10B *Escherichia coli* cells (GIBCO). Cosmids were organized into a contig by PCR, hybridization, and restriction pattern analyses and transferred to *Agrobacterium tumefaciens* strain GV3101 by conjugation in a triparental mating with helper strain HB101/PRK2013.

After identification of a complementation interval, the cosmid 2–63 was digested with *Hind*III and self-ligated, generating cosmid 2–63H. The cosmid 2–8 was digested completely with *Xba*I and partially with *Bsa*AI and self-ligated, resulting in cosmid 2–8XB. Cosmids 2–63H and 2–8XB were transferred to *A. tumefaciens* as described above.

Transformation of *Arabidopsis*. *A. tumefaciens* strains containing cosmids were grown under selection in Luria broth for 24–48 hr at 30°C. Cells were harvested and resuspended in infiltration medium (10 mM MgCl₂, 5% sucrose, 0.05% Silwet L-77, and 44 nM benzylaminopurine) to an OD₆₀₀ of 0.8. C24 plants (approximately 30 plants per clone) were transformed by vacuum infiltration (16). Seeds from infiltrated plants were surface-sterilized and plated on selective medium (0.5× Murashige and Skoog MS medium, 1× B5 vitamins, and 25 µg/ml GA).

Growth, Inoculation, and β-Glucuronidase (GUS) Activity Assays of Transformants. Transformants were transferred to soil and grown for 15–20 days before inoculation. Rosette leaves were dusted with carborundum and inoculated with TEV-GUS, a recombinant TEV strain encoding GUS (17). GUS activity assays were done by using inflorescence tissue at 15 and 20 days postinoculation (11). Inoculated leaves from plants that scored negative in GUS activity assays of systemic tissue were tested for infection by *in situ* histochemical assay (17).

Analysis of DNA and Protein Sequences. Sequencing reactions using a dye terminator cycle sequencing protocol (Applied Biosystems) were done in a Robocycler 96 thermocycler (Stratagene), analyzed on an Applied Biosystems 373 or 377 automated sequencer, and compiled by using the GCG program (Wisconsin Package Version 10.0). Similarity searches of sequence databases were done by using BLAST (18). A phylogenetic tree was constructed by using a heuristic search and bootstrap analysis with PAUPSEARCH and PAUPDISPLAY (Wisconsin Package Version 10.0, GCG).

Isolation and Analysis of the *RTM1* cDNA. Total RNA was isolated from rosette leaves of Col-0 plants and used to prepare poly(A)⁺ RNA by using an oligo(dT)-coated latex resin (Qiagen, Chatsworth, CA). Complementary DNA was synthesized by using 250 ng of poly(A)⁺ RNA and was subjected to 5′ and 3′ rapid amplification of cDNA ends (RACE) using the SMART kit (CLONTECH). The RACE products were gel-purified and sequenced. Primers containing *Xho*I and *Bam*HI sites were designed at the 5′ and 3′ ends of RACE products and used to amplify the full-length *RTM1* cDNA. The full-length cDNA was transferred to the pRTL2 vector (19), which contains a cauliflower mosaic virus 35S promoter and 35S terminator. The

expression cassette was excised and subcloned into pSLJ75515. This construct was mobilized into *A. tumefaciens* and used to transform C24 plants as described above.

Results

Mapping of *RTM1*. Screens of ethyl methanesulfonate-mutagenized *Arabidopsis* Col-0 populations previously revealed several mutants with gain of susceptibility or loss of resistance to TEV (12). These mutants, many of which contained mutations at the *RTM1* locus, allow long-distance movement of TEV in an ecotype that normally restricts the virus to inoculated rosette leaves. An *RTM1* mapping population was generated by crossing the A133 mutant (*rtm1-2* allele) to the nonsusceptible (*RTM1*) ecotype Ws-2. Approximately 7,000 F₂ individuals were inoculated with TEV-bar, a recombinant TEV strain encoding the bacterial *bar* gene that confers resistance to the herbicide GA (12). Plants containing a *rtm1-2/rtm1-2* genotype were susceptible to systemic infection by TEV-bar, whereas *RTM1/rtm1-2* or *RTM1/RTM1* plants restricted TEV-bar to inoculated leaves. Systemically infected plants were insensitive to GA and survived herbicide application. Tissue was recovered from 590 of the most vigorous GA survivors.

Analysis of recombination events between molecular markers and the *RTM1* locus in susceptible F₂ individuals confirmed that *RTM1* was located between markers ATEAT1 and F17J18-M on chromosome 1. Additional polymorphic CAPS and dCAPS markers were identified within this interval by using sequence from yeast artificial chromosome and BAC ends 8H12R, F2M5-T7, T12A12-Sp6, and F3F20-Sp6. Of 1,120 chromosomes examined, three recombination events were identified between *RTM1* and 8H12R (Fig. 1A). Markers T12A12-Sp6, F3F20-Sp6, and F2M5-T7 cosegregated with *RTM1*. There were five recombination events between *RTM1* and F17J18-M of 1,180 chromosomes (Fig. 1A). We concluded that *RTM1* was located between markers 8H12R and F17J18-M.

Complementation Cloning of *rtm1*. A BAC contig spanning the region between ATEAT1 and F17J18-M was assembled (Fig. 1A). BACs F26O9, F25G24, and F2M5 were fragmented and subcloned into the binary cosmid vector pSLJ75515 to generate a cosmid contig (Fig. 1A). Cosmids were introduced into the *Arabidopsis* C24 genome (*rtm1/rtm1*) by *Agrobacterium*-mediated transformation, and transformants were inoculated with TEV-GUS and examined for systemic infection at approximately 20 days postinoculation. Cosmids containing the *RTM1* allele were expected to confer an infection phenotype in which virus was restricted to inoculated leaves. The majority of plants (T1 generation) containing six of the cosmids, each overlapping a common sequence of approximately 5 kb, complemented the *rtm1* phenotype of C24 (Fig. 1A). The TEV-restrictive phenotype of these plants was inherited in the T2 generation (Fig. 1B and data not shown). None of the cosmids that lacked this 5-kb region conferred a nonsusceptible phenotype. The overlapping interval was sequenced and three putative ORFs were identified. To determine which portion of the 5-kb region contained *RTM1*, cosmid clones containing portions of the complementing interval were generated and transferred into C24 (Fig. 1A). The clone 2–8XB contained all three ORFs from the complementing interval, but no sequence from outside this region. The cosmid 2–63H contained only the most centromeric ORF (ORF3) from the complementing interval. Both 2–8XB and 2–63H conferred a restricted movement phenotype to C24 plants (Fig. 1A), suggesting that ORF3 and surrounding sequence was sufficient to complement the *rtm1* phenotype.

The approximately 5-kb complementing interval containing the three putative ORFs from mutants A130, A133, and A165 was sequenced. Comparison of these sequences with wild-type Col-0 revealed a single nucleotide difference in each mutant

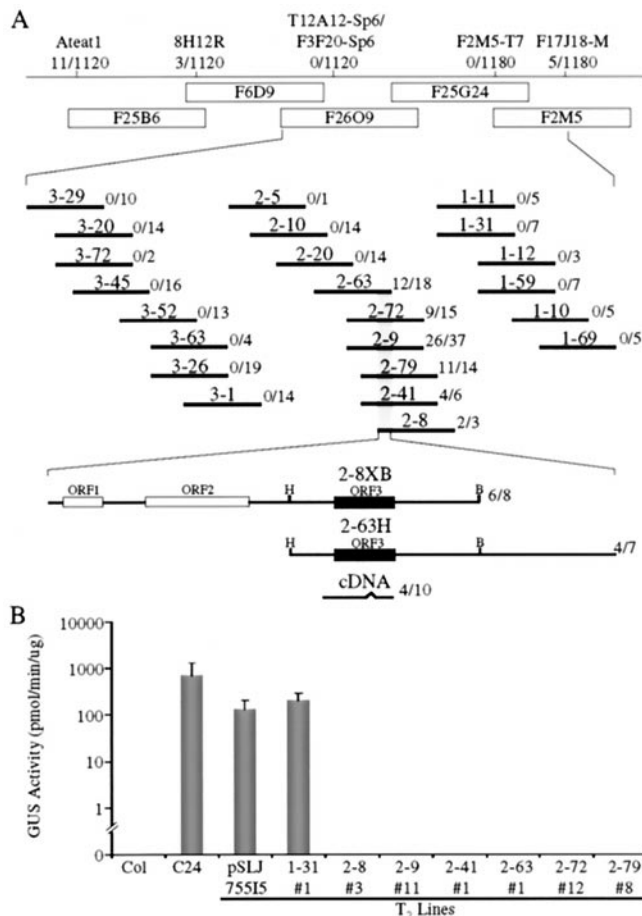


Fig. 1. Cloning of *RTM1*. (A) PCR-based markers that flank or cosegregate with *RTM1* are indicated at the top. The number of recombination events per total number of meiotic events scored is given below each marker. A BAC contig spanning *RTM1* is shown (open boxes). Cosmids (solid lines) derived from BACs F2M5 (cosmids with prefix 1), F25G24 (prefix 2), and F26O9 (prefix 3) were introduced into *Arabidopsis* ecotype C24 (*rtm1/rtm1*). The complementing interval is shaded. Putative ORFs are indicated by boxes in the bottom expanded region. The *Hind*III (H) and *Bsa*AI (B) restriction sites used to generate 2–63H and 2–8XB, respectively, are indicated. The number of putative C24 transformants that restricted long-distance movement of TEV-GUS per total plants tested is shown adjacent to each clone. (B) GUS activity assays of selected TEV-GUS-infected T2 C24 lines containing complementing cosmids (2–8 #3, 2–9 #11, 2–41 #1, 2–63 #1, 2–72 #12, 2–79 #8), a noncomplementing cosmid (1–31 #1), or empty vector (pSLJ75515), or wild-type susceptible (C24) and nonsusceptible (Col-0) lines. Inflorescence tissue from 10 T2 individuals was tested at 20 days postinoculation. The mean GUS activity value (+ SD) is shown. (C) Representation of the *RTM1* cDNA, with the nucleotide positions of the start codon (nucleotide 55), 5' intron splice site (nucleotide 262), stop codon (nucleotide 577), and 3' terminal nucleotide (nucleotide 729) indicated. Arrows indicate positions of amino acid substitutions in *rtm1* mutant alleles.

within ORF3 (see below). No other differences between the mutant and wild-type sequences were detected. Therefore, ORF3 was concluded to encode *RTM1*.

Using PCR primers based on the genomic sequence, the ends of the *RTM1* mRNA were identified by RACE. The RACE products were sequenced and primers were designed to amplify the full-length cDNA. The *RTM1* cDNA is 729 nt in length and contains a 119-nt intron (Fig. 1C). The deduced *RTM1* protein is 174 aa in length, assuming that translation initiates at the first

Col	MKIGFVGKHDARSTTIVNWDGSHDGFISQIFLSHGAVAGIMSIQF	45
C24	MKIGFVGKHDARSTTIVNWDGSHDGFISQIFLSHGAVAGIMSIQF	45
La- <i>er</i>	MKIGFVGKHDARSTTIVNWDGSHDGFISQIFLSHGAVAGIMSIQF	45
<i>rtm1-1</i>	MKIGFVGKHDARSTTIVNWDGSHDGFISQIFLSHGAVAGIMSIQF	45
<i>rtm1-2</i>	MKIGFVGKHDARSTTIVNWDGSHDGFISQIFLSHGAVAGIMSIQF	45
<i>rtm1-5</i>	MKIGFVGKHDARSTTIVNWDGSHDGFISQIFLSHGAVAGIMSIQF	45
Col	QFVMDGKLVLSDRHGPFSGNMFDVIELNYPHEYITGISGEYKYKYE	90
C24	QFVMDGKLVLSDRHGPFSGNMFDVIELNYPHEYITGISGEYKYKYE	90
La- <i>er</i>	QFVMDGKLVLSDRHGPFSGNMFDVIELNYPHEYITGISGEYKYKYE	90
<i>rtm1-1</i>	QFVMDGKLVLSDRHGPFSGNMFDVIELNYPHEYITGISGEYKYKYE	90
<i>rtm1-2</i>	QFVMDGKLVLSDRHGPFSGNMFDVIELNYPHEYITGISGEYKYKYE	90
<i>rtm1-5</i>	QFVMDGKLVLSDRHGPFSGNMFDVIELNYPHEYITGISGEYKYKYE	90
Col	ANNPHMRSCLKFNTNTSEYGPFGTSGSSNDKFAFKLGKSPQFGGFH	135
C24	ANNPHMRSCLKFNTNTSEYGPFGTSGSSNDKFAFKLGKSPQFGGFH	135
La- <i>er</i>	ANNPHMRSCLKFNTNTSEYGPFGTSGSSNDKFAFKLGKSPQFGGFH	135
<i>rtm1-1</i>	ANNPHMRSCLKFNTNTSEYGPFGTSGSSNDKFAFKLGKSPQFGGFH	135
<i>rtm1-2</i>	ANNPHMRSCLKFNTNTSEYGPFGTSGSSNDKFAFKLGKSPQFGGFH	135
<i>rtm1-5</i>	ANNPHMRSCLKFNTNTSEYGPFGTSGSSNDKFAFKLGKSPQFGGFH	135
Col	GTIDASGLQYIGVYLRPKTVLPKIDTGNAEETESKIVLG	174
C24	GTIDASGLQYIGVYLRPKTVLPKIDTGNAEETESKIVLG	174
La- <i>er</i>	GTIDASGLQYIGVYLRPKTVLPKIDTGNAEETE-----	168
<i>rtm1-1</i>	GTIDASGLQYIGVYLRPKTVLPKIDTGNAEETESKIVLG	174
<i>rtm1-2</i>	GTIDASGLQYIGVYLRPKTVLPKIDTGNAEETESKIVLG	174
<i>rtm1-5</i>	GTIDASGLQYIGVYLRPKTVLPKIDTGNAEETESKIVLG	174

Fig. 2. Predicted *RTM1* amino acid sequence for Col-0, C24, and La-*er* alleles, and mutant alleles *rtm1-1*, *rtm1-2*, and *rtm1-5*. Differences relative to the Col-0 sequence are shaded. Dashes indicate amino acid residues absent from the deduced La-*er* product.

AUG codon. The *RTM1* cDNA was cloned into an intermediate vector containing the cauliflower mosaic virus 35S promoter and terminator sequences. This expression cassette was transferred to the cosmid vector, and transgenic *Arabidopsis* C24 plants were produced. The 35S-*RTM1* cDNA restricted long-distance movement in four of 10 transgenic plants (Fig. 1A).

Hypothetical translation products of the mutant alleles *rtm1-1*, *rtm1-2*, and *rtm1-5* (A130, A133, and A165, respectively), as well as the *rtm1* alleles of C24 and La-*er*, were examined (Fig. 2). Single nucleotide changes in the mutants resulted in substitutions of asparagine for aspartate 139, phenylalanine for serine 56, and aspartate for glycine 132, respectively (Fig. 2). The C24 *rtm1* allele contained four codon differences relative to Col-0, including tyrosine for serine 29, cysteine for phenylalanine 62, aspartate for asparagine 65, and isoleucine for asparagine 93, as well as a 26-nt deletion within the intron (Fig. 2, data not shown). The La-*er* *rtm1* allele contained a stop codon rather than a codon for serine 169, resulting in a protein lacking six C-terminal residues relative to the Col-0 *RTM1* protein.

Similarity of *RTM1* to Jacalin and Jacalin-Related Proteins. Database searches revealed similarity (23% identical and 33% similar at the amino acid level over the length of *RTM1*) between the predicted *RTM1* protein and the polypeptide chains of jacalin (20), a D-galactose-specific lectin from *Artocarpus integrifolia* (Fig. 3 and Table 1). Jacalin is synthesized as a prepropeptide that is proteolytically processed to form two mature polypeptides, a 133-aa α -chain and a 20- to 21-aa β -chain (21), which remain associated in the folded protein (22). The majority of *RTM1* sequence is similar to the jacalin α -chain, although the N terminus of *RTM1* exhibits similarity with the C-terminal half of the β -chain (Fig. 3). The C-terminal 22 residues of *RTM1* lack similarity to either jacalin chain. Four jacalin α -chain residues involved in carbohydrate binding (22)—Gly-1, Tyr-122, Trp-123, and Asp-125—are not conserved in *RTM1*. The *RTM1* sequence also bears similarity to the

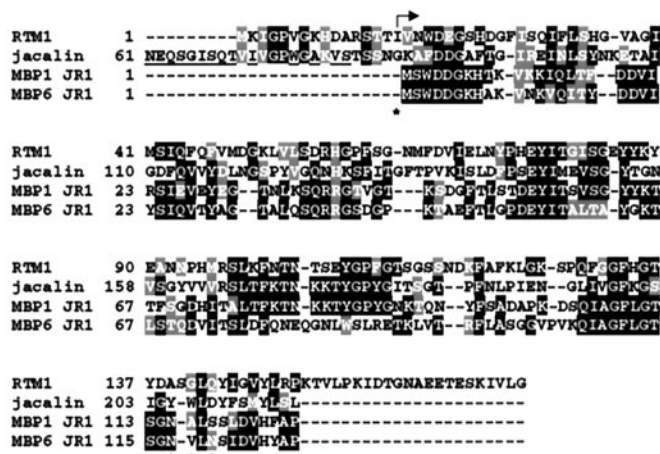


Fig. 3. Alignment of the deduced Col-0 RTM1 sequence, amino acids 61–217 of the jacalin prepropeptide, and JRs from *B. napus* MBPs 1 and 6. Gray and black shading indicate conserved and identical residues, respectively. The arrow indicates the first residue in the jacalin α -chain. Amino acids corresponding to the mature jacalin β -chain are underlined. * indicate residues involved in carbohydrate binding in the jacalin α polypeptide (22). Dashes indicate gaps in the alignment.

sequences of jacalin-related lectins, such as the *Maclura pomifera* agglutinin (20) (Fig. 3 and Table 1).

Database searches also revealed convincing similarity between RTM1 and a number of proteins from plants in the *Brassicaceae* family (Fig. 3 and Table 1). Six of these proteins from *Brassica napus* were shown previously to be myrosinase-binding proteins (MBPs) (23–25). These proteins copurify in a

bound form with myrosinase, an enzyme that catalyzes hydrolysis of glucosinolates. Hydrolysis of glucosinolates occurs upon wounding or tissue damage and results in release of volatile, toxic compounds (for example, thiocyanates and isothiocyanates) with insecticidal and antifungal properties (26). In fact, expression of these MBP genes was shown previously to be induced by wounding and treatment with methyl jasmonate (23–25). In addition to MBPs from *B. napus*, nine inferred *Arabidopsis* proteins were identified as proteins containing jacalin-like domains. Two of these were isolated as cDNA sequences of genes (termed f-AtMBP and A. th. floral cDNA) expressed preferentially in inflorescence tissue. Seven proteins were identified from genes inferred from *Arabidopsis* genome sequences. Interestingly, all but two of these *B. napus* and *Arabidopsis* proteins were found to contain multiple repeats of the RTM1-like sequence, most of which corresponded precisely to the jacalin α -chain (Fig. 3 and 4). For this reason, the repeated unit was designated as the jacalin repeat (JR). In proteins containing multiple JRs, between two and six copies were identified (Fig. 4). In a few proteins, partial homologues were identified, and in one case, the JR sequence was found to be fused to another domain comprised of “kelch repeats,” which likely interact with actin microfilaments (27).

The JR Domain Family. Fifty-five complete JRs from *B. napus* and *Arabidopsis* were identified in database searches. The phylogenetic relationship of RTM1 to other JR-containing proteins was inferred from cladistic analysis using individual JR domains from each protein. Surprisingly, RTM1 grouped in a clade with JR1 from four *B. napus* MBPs (Fig. 5, color-coded red). No other *Arabidopsis* or *B. napus* JR domains were identified in this group. This grouping was reinforced through generation of a phylogram using distance analysis (data not shown). Several other clades

Table 1. Functions of proteins containing JRs

Name	Species	Accession number	JRs*	Function	Ref.
RTM1	<i>A. th.</i>	AF191302	1	Restricts long-distance movement of TEV	
Jacalin α -chain	<i>A. in.</i>	P18670	1	Seed lectin; specific for Gal β 1-3GalNAc; binds human IgA1	20
MPA	<i>M. p.</i>	P18674	1	Seed lectin; specific for Gal β 1-3GalNAc	20
f-AtMBP	<i>A. th.</i>	BAA82151	4	Specifically expressed in flowers	
Floral cDNA	<i>A. th.</i>	BAA22099	3	Preferentially expressed in flowers	28
F25118-19	<i>A. th.</i>	AAC04913	1	Hypothetical protein (chromosome II)	
T204-6	<i>A. th.</i>	AAB63633	2	Hypothetical protein (chromosome III)	
F11P17-5	<i>A. th.</i>	AAB71472	4	Hypothetical protein (chromosome I)	
T13D8-1	<i>A. th.</i>	AAC24045	3	Hypothetical protein (chromosome I)	
T19L18-21	<i>A. th.</i>	AAC31237	3	Hypothetical protein (chromosome II)	
F5F19-17	<i>A. th.</i>	AAD12685	4	Hypothetical protein (chromosome I)	
T16B24-3	<i>A. th.</i>	AAC28977	3	Hypothetical protein (chromosome II)	
MBP 1	<i>B. n.</i>	CAA70587	6	Binds myrosinase; JA- and wound-inducible; implicated in defense	24
MBP 2	<i>B. n.</i>	AAC08048	5	Binds myrosinase; JA- and wound-inducible; implicated in defense	25
MBP 3	<i>B. n.</i>	AAC08049	3	Binds myrosinase; JA- and wound-inducible; implicated in defense	25
MBP 4	<i>B. n.</i>	AAC08050	1	Binds myrosinase; JA- and wound-inducible; implicated in defense	25
MBP 5	<i>B. n.</i>	CAA72270	5	Binds myrosinase; JA- and wound-inducible; implicated in defense	23
MBP 6	<i>B. n.</i>	CAA72271	4	Binds myrosinase; JA- and wound-inducible; implicated in defense	23
MBP homolog	<i>B. n.</i>	AAC08051	3	JA- and wound-inducible	25

A. th., *A. thaliana*; *A. in.*, *A. integrifolia*; *M. p.*, *M. pomifera*; *B. n.*, *B. napus*.

*Number of complete JRs.

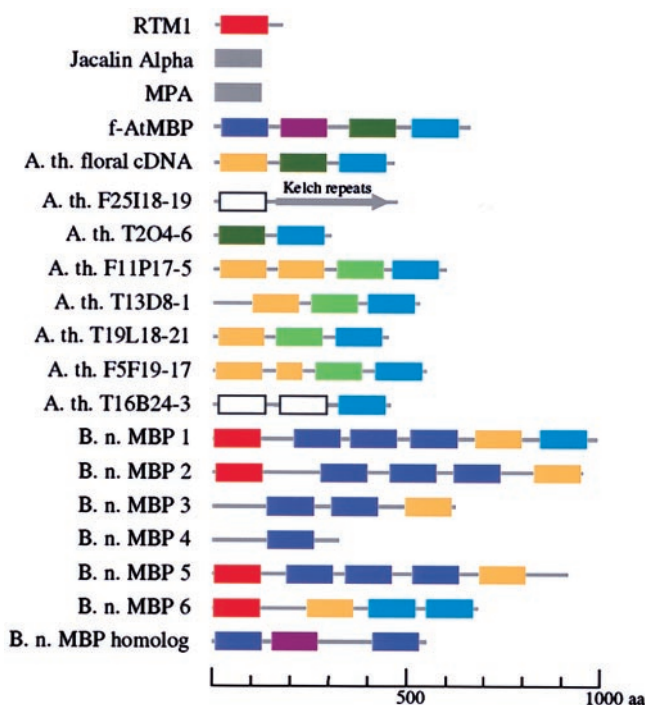


Fig. 4. Organization of proteins containing JRs. Boxes indicate JRs. Color of boxes corresponds to the position of the repeat in the cladogram (Fig. 5). Open boxes indicate repeats that are not grouped within a clade with other JRs in Fig. 5. Gray arrow indicates the Kelch repeats of *A. th.* F25I18-19. A. in., *A. integrifolia*; A. th., *A. thaliana*; B. n., *B. napus*; f-, floral; M. p., *M. pomifera*.

were identified. For example, most, but not all, of the JRs from *B. napus* MBPs grouped within one major clade (Fig. 5, color-coded purple). Another major clade, which was supported in 100% of the bootstrap replicates, contained JRs from several *B. napus* MBPs and *Arabidopsis* inferred proteins (Fig. 5, color-coded orange). Three JRs, each from *Arabidopsis* inferred proteins, did not fall into any group (Fig. 5).

Evidence consistent with JR domain extragenic duplication and shuffling, intragenic domain duplication, and whole gene duplication was obtained in the phylogenetic analysis (color codes in Fig. 4 indicate phylogenetic relationships inferred in Fig. 5). Extragenic duplication appears to have occurred on multiple occasions, as unique combinations of closely or distantly related JRs were identified in several proteins. For example, *Arabidopsis* f-AtMBP JR3 and JR4 are closely related to *Arabidopsis* floral cDNA protein JR2 and JR3, respectively, although f-AtMBP JR1 and JR2 and floral cDNA protein JR1 are from different clades (Figs. 4 and 5). Intragenic JR domain duplication may explain why JR2, JR3, and JR4 of *B. napus* MBP 1 are so closely related (Figs. 4 and 5). In addition, whole gene or multidomain duplication may explain why clusters of JRs are conserved among several proteins. This is seen clearly with JR1-JR5 from both *B. napus* MBP1 and MBP2, where the corresponding JRs were always grouped in the tree (Figs. 4 and 5). This phylogenetic analysis suggests that RTM1 is in a domain family with members that have duplicated, rearranged in novel combinations, and diversified.

Discussion

The *RTM1* gene encodes a protein similar to the lectin jacalin (20) and a large family of related proteins containing one or more repeats of a jacalin-like domain. This finding was unexpected, as jacalin-related proteins have been shown or suggested to be involved in resistance against insects and fungi (26, 29) but

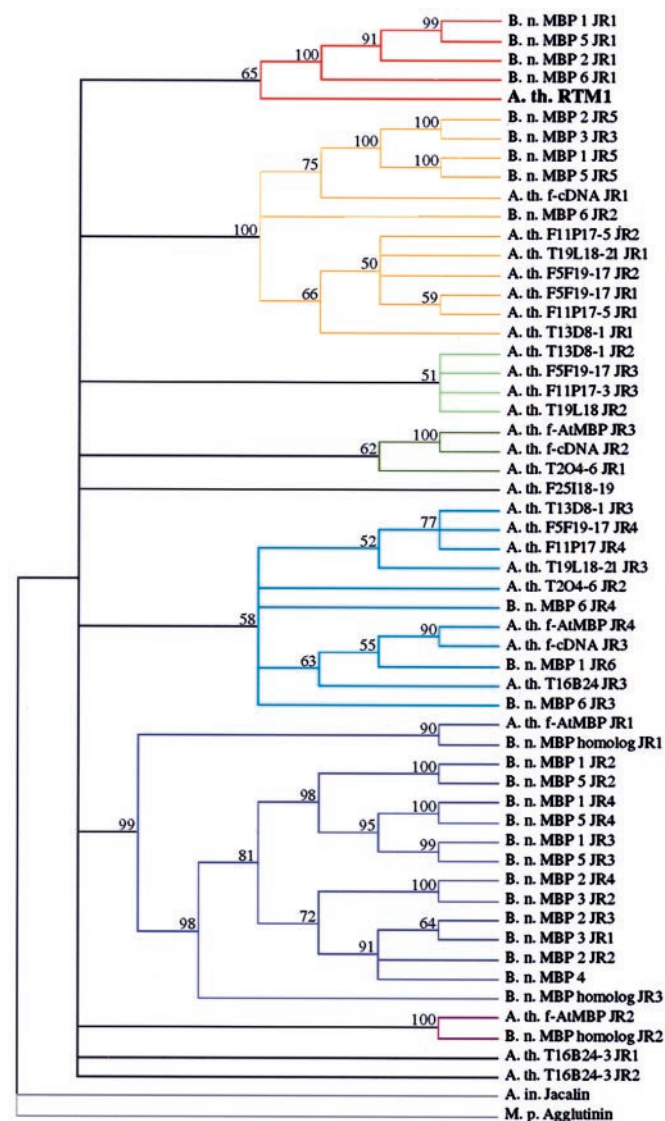


Fig. 5. Cladistic analysis of JRs from various proteins. Repeats were aligned and assembled into a single tree by a heuristic search using bootstrap analysis (100 replicates) with simple sequence addition, tree-bisection-reconnection branch swapping, steepest descent off, and MULPARS on using PAUPSEARCH and PAUPDISPLAY. Jacalin and *M. pomifera* agglutinin were assigned to the out-group. Numbers above branches are bootstrap percentage values. The relative positions of JRs in proteins containing multiple repeats are indicated (e.g., JR1, JR2). Colored lines highlight each clade. For abbreviations, see Fig. 4 legend.

by mechanisms that would seem to be unrelated to virus infection. The *M. pomifera* agglutinin, a lectin with a single JR, inhibits development of the cowpea weevil (29). Indeed, lectins in general have been shown to confer defensive properties against bacteria, fungi, and insects (30). The *B. napus* MBPs, most of which contain multiple JRs, are induced by wounding/jasmonic acid (JA) and interact with myrosinase (23-25). Glucosinolate accumulation also is induced in wounded or JA-treated plants (26, 31). This finding leads to the hypothesis that MBPs are up-regulated systemically in response to JA hormone signaling during insect or fungal attack. Whether or not up-regulation of MBPs affects the activity of myrosinase, thereby leading to more efficient or controlled glucosinolate hydrolysis, remains to be determined.

The restricted movement phenotype conditioned by *RTM1* is specifically active against TEV. The Col-0 ecotype (*RTM1*) is

fully susceptible to a range of other viruses, such as turnip crinkle virus, turnip mosaic virus, and cauliflower mosaic virus (5, 32, 33). Does the *RTM1* gene confer TEV specificity, or does RTM1 participate in a complex or pathway in which specificity is determined by another component? In cases of resistance involving NBS-LRR-type *R* genes, natural variation in resistance almost always is caused by variation at the *R* gene locus. A feature in common between RTM1 and NBS-LRR-type *R* gene products is the abundance of genes, in a given plant, encoding similar proteins. This feature could indicate that the JR, like NBS-LRR proteins, is particularly suitable for adapting to unique functions requiring different specificities. The structures of MBPs and several inferred *Arabidopsis* proteins further suggest that JRs are modular units that can be duplicated and combined in ways to provide unique functions.

The lack of similarity between RTM1 and components of gene-for-gene-type resistance suggests that resistance to TEV in *Arabidopsis* may occur by a fundamentally different process. In fact, RTM1-mediated resistance involves neither a hypersensitive reaction nor systemic acquired resistance, which typically are associated with gene-for-gene-type resistance. Two hypotheses for how RTM1 mediates TEV-specific resistance are viable. First, RTM1, in cooperation with other factors, could block long-distance movement of TEV by preventing entry into, passage through, or exit from the phloem. This could occur if RTM1, or a RTM1-containing complex, physically interacts with one or more of the TEV factors required for long-distance movement (HC-Pro, CI, NIa, and capsid) (34–39). The RTM1

protein or complex also could inhibit the function of cellular factors or structures required for long-distance movement. The lectin-like structure of RTM1 suggests this protein may interact with a glycoconjugate, although carbohydrate-binding activities of RTM1 have not been investigated. Second, RTM1 could participate as a structural or signaling component required for establishing a TEV-restrictive state in systemic, noninoculated tissue. In this scenario, RTM1 could potentially be involved in generation, perception, or transport of a systemic signal, or be required to establish the TEV-restrictive state in distal cells. Either of these possibilities would account for the infection phenotype of TEV in Col-0, namely, that virus is absent from tissues distal to the initial infection sites but is not impeded in inoculated leaves. Further insight into the RTM1-mediated restriction of TEV will come from cloning and analysis of *RTM2* and *RTM3*, as well as from identification and analysis of intergenic suppressor mutants.

Finally, it remains to be determined the extent to which resistance to other viruses in *Arabidopsis* and other plants is conditioned by mechanisms similar to those involving RTM1. It seems unlikely that the TEV-*Arabidopsis* restriction system involving RTM1 represents an isolated example.

We thank Jonathan Jones for providing the plasmid pSLJ75515. We are grateful to Pam Soltis for helpful discussion regarding phylogenetic analysis and Derek Pouchnik for invaluable sequencing assistance. This work was supported by grants from the National Institutes of Health (AI43288) and the United States Department of Agriculture (98-35303-6485).

1. Staskiewicz, B. J., Ausubel, F. M., Baker, B. J., Ellis, J. G. & Jones, J. D. (1995) *Science* **268**, 661–667.
2. Whitham, S., Dinesh-Kumar, S. P., Choi, D., Hehl, R., Corr, C. & Baker, B. (1994) *Cell* **78**, 1101–1115.
3. Bendahmane, A., Kanyuka, K. & Baulcombe, D. C. (1999) *Plant Cell* **11**, 781–792.
4. Hammond-Kosack, K. E. & Jones, J. D. G. (1997) *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**, 575–607.
5. Callaway, A., Liu, W., Andrianov, V., Stenzler, L., Zhao, J., Wettlaufer, S., Jayakumar, P. & Howell, S. H. (1996) *Mol. Plant Microbe Interact.* **9**, 810–818.
6. Nicolas, O., Dunnington, S. W., Gotow, L. F., Pirone, T. P. & Hellmann, G. M. (1997) *Virology* **237**, 452–459.
7. Schaad, M. C. & Carrington, J. C. (1996) *J. Virol.* **70**, 2556–2561.
8. Ratcliff, F., Harrison, B. D. & Baulcombe, D. C. (1997) *Science* **276**, 1558–1560.
9. Al-Kaff, N. S., Covey, S. N., Kreike, M. M., Page, A. M., Pinder, R. & Dale, P. J. (1998) *Science* **279**, 2113–2115.
10. Kooter, J. M., Matzke, M. A. & Meyer, P. (1999) *Trends Plant Sci.* **4**, 340–347.
11. Mahajan, S. K., Chisholm, S. T., Whitham, S. & Carrington, J. C. (1998) *Plant J.* **14**, 177–186.
12. Whitham, S., Yamamoto, M. & Carrington, J. C. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 772–777.
13. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680.
14. Marck, C. (1988) *Nucleic Acids Res.* **16**, 1829–1836.
15. Jones, J. D., Shlumukov, L., Carland, F., English, J., Scofield, S. R., Bishop, G. J. & Harrison, K. (1992) *Transgenic Res.* **1**, 285–297.
16. Bechtold, N. & Pelletier, G. (1998) *Methods Mol. Biol.* **82**, 259–266.
17. Dolja, V. V., McBride, H. J. & Carrington, J. C. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10208–10212.
18. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990) *J. Mol. Biol.* **215**, 403–410.
19. Restrepo, M. A., Freed, D. D. & Carrington, J. C. (1990) *Plant Cell* **2**, 987–998.
20. Young, N. M., Johnston, R. A. & Watson, D. C. (1991) *FEBS Lett.* **191**, 382–384.
21. Kabir, S. (1998) *J. Immunol. Methods* **212**, 193–211.
22. Sankaranarayanan, R., Sekar, K., Banerjee, R., Sharma, V., Surolia, A. & Vijayan, M. (1996) *Nat. Struct. Biol.* **3**, 596–603.
23. Geshi, N. & Brandt, A. (1998) *Planta* **204**, 295–304.
24. Taipalensuu, J., Eriksson, S. & Rask, L. (1997) *Eur. J. Biochem.* **250**, 680–688.
25. Taipalensuu, J., Falk, A., Ek, B. & Rask, L. (1997) *Eur. J. Biochem.* **243**, 605–611.
26. Bones, A. M. & Rossiter, J. (1996) *Physiol. Plant* **97**, 194–208.
27. Robinson, D. N. & Cooley, L. (1997) *J. Cell Biol.* **138**, 799–810.
28. Utsugi, S., Sakamoto, W., Ogura, Y., Murata, M. & Motoyoshi, F. (1996) *Plant Mol. Biol.* **32**, 759–765.
29. Murdock, L. L., Huesing, J. E., Nielsen, S. S., Pratt, R. C. & Shade, R. E. (1990) *Phytochemistry* **29**, 85–89.
30. Peumans, W. J. & Damme, E. J. M. v. (1995) *Histochem. J.* **27**, 253–271.
31. Bodnaryk, R. P. (1994) *Phytochemistry* **35**, 301–305.
32. Simon, A. E., Li, X. H., Lew, J. E., Stange, R., Zhang, C., Polacco, M. & Carpenter, C. D. (1992) *Mol. Plant-Microbe Interact.* **5**, 496–503.
33. Sanchez, F., Martinez-Herrera, D., Aguilar, I. & Ponz, F. (1998) *Virus Res.* **55**, 207–219.
34. Dolja, V. V., Haldeman, R., Robertson, N. L., Dougherty, W. G. & Carrington, J. C. (1994) *EMBO J.* **13**, 1482–1491.
35. Dolja, V. V., Haldeman-Cahill, R., Montgomery, A. E., VandenBosch, K. A. & Carrington, J. C. (1995) *Virology* **207**, 1007–1016.
36. Cronin, S., Verchot, J., Haldeman-Cahill, R., Schaad, M. C. & Carrington, J. C. (1995) *Plant Cell* **7**, 549–559.
37. Kasschau, K. D., Cronin, S. & Carrington, J. C. (1997) *Virology* **228**, 251–262.
38. Schaad, M. C., Lellis, A. D. & Carrington, J. C. (1997) *J. Virol.* **71**, 8624–8631.
39. Carrington, J. C., Jensen, P. E. & Schaad, M. C. (1998) *Plant J.* **14**, 393–400.