The Plant Cell, Vol. 10, 1817–1832, November 1998, www.plantcell.org © 1998 American Society of Plant Physiologists

The Major Resistance Gene Cluster in Lettuce Is Highly Duplicated and Spans Several Megabases

Blake C. Meyers,[1] Doris B. Chin, Katherine A. Shen, Subramoniam Sivaramakrishnan,[2] Dean O. Lavelle, Zhen Zhang,[3] and Richard W. Michelmore[4]

Department of Vegetable Crops, University of California, Davis, California 95616

At least 10 *Dm* genes conferring resistance to the oomycete downy mildew fungus *Bremia lactucae* map to the major resistance cluster in lettuce. We investigated the structure of this cluster in the lettuce cultivar Diana, which contains *Dm3.* A deletion breakpoint map of the chromosomal region flanking *Dm3* was saturated with a variety of molecular markers. Several of these markers are components of a family of resistance gene candidates (*RGC2*) that encode a nucleotide binding site and a leucine-rich repeat region. These motifs are characteristic of plant disease resistance genes. Bacterial artificial chromosome clones were identified by using duplicated restriction fragment length polymorphism markers from the region, including the nucleotide binding site–encoding region of *RGC2.* Twenty-two distinct members of the *RGC2* family were characterized from the bacterial artificial chromosomes; at least two additional family members exist. The *RGC2* family is highly divergent; the nucleotide identity was as low as 53% between the most distantly related copies. These *RGC2* genes span at least 3.5 Mb. Eighteen members were mapped on the deletion breakpoint map. A comparison between the phylogenetic and physical relationships of these sequences demonstrated that closely related copies are physically separated from one another and indicated that complex rearrangements have shaped this region. Analysis of low-copy genomic sequences detected no genes, including *RGC2*, in the *Dm3* region, other than sequences related to retrotransposons and transposable elements. The related but divergent family of *RGC2* genes may act as a resource for the generation of new resistance phenotypes through infrequent recombination or unequal crossing over.

INTRODUCTION

Disease resistance genes frequently occur in tightly linked clusters (Pryor, 1987; Crute and Pink, 1996; Michelmore and Meyers, 1998). Clusters of plant resistance genes were first established by use of classic genetic techniques; detailed molecular analyses are now beginning to unravel the complexity of these loci and the underlying mechanisms determining their structure (Parniske et al., 1997; Song et al., 1997). It is becoming increasingly apparent that such clusters may be both common and complex genomic regions in plants.

Clusters of resistance genes have been identified in diverse plant species. More than 30 different resistance specificities to the single fungal pathogen responsible for flax rust disease, *Melampsora lini*, have been mapped to five linkage groups (Flor, 1971; Islam and Shepherd, 1991). These loci exemplify two possible genetic arrangements that may exist for clusters of resistance genes: the flax *L* locus contains at least 13 allelic rust resistance specificities, and the more complex *M* locus exists as a tandem array of at least seven genes (Islam and Shepherd, 1991). In maize, multiple *Rp* genes, both linked and allelic, have been observed to mediate resistance to the rust fungus *Puccinia sorghi*; 16 genetically separable loci were mapped to a single cluster known as the *Rp1* complex (Saxena and Hooker, 1968; Hulbert, 1997). Complex disease resistance clusters also have been identified in lettuce (Farrara et al., 1987; Witsenboer et al., 1995), Arabidopsis (Kunkel, 1996; Holub, 1997), rice (Song et al., 1995), barley (Jorgensen, 1994), tomato (Jones et al., 1993), and other plant species (reviewed in Michelmore and Meyers, 1998). Several of the specificities within these genetically well-defined resistance loci have been targeted for molecular cloning and analysis.

The molecular characterization of resistance gene clusters has been advanced by the recent cloning of plant disease resistance genes from diverse species. Most of these genes contain regions encoding leucine-rich repeats (LRRs), with

[1] Current address: DuPont Agricultural Biotechnology, Delaware Technology Park, Newark, DE 19714.
[2] Current address: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) Asia Center, Patacheru 502 324, Andhra Pradesh, India.
[3] Current address: Institute of Molecular Agrobiology, National University of Singapore, 59A The Fleming, 1 Science Park Dr., Singapore.
[4] To whom correspondence should be addressed. E-mail rwmichelmore @ucdavis.edu; fax 530-752-9659.

or without either a nucleotide binding site (NBS) or a kinase domain (reviewed in Baker et al., 1997; Hammond-Kosack and Jones, 1997). These structurally similar gene products exhibit dramatic differences in specificity and can confer resistance to fungal, viral, nematode, or bacterial pathogens. Many, but not all, resistance genes are members of multigene families. DNA gel blot hybridization and genetic analysis indicate that the *Pto* and *Cf* genes of tomato, the *Xa21* gene of rice, and the *N* gene of tobacco all exist as tandem arrays of sequences (Martin et al., 1993; Whitham et al., 1994; Song et al., 1995; Dixon et al., 1996; Parniske et al., 1997). Sequences similar to resistance genes are being isolated via degenerate oligonucleotides or identified from cDNA libraries and are often linked to known resistance loci (Kanazin et al., 1996; Leister et al., 1996; Yu et al., 1996; Botella et al., 1997; Lagudah et al., 1997; Aarts et al., 1998). These resistance gene candidates also exist as multigene families, although their function has yet to be demonstrated. The multigenic structure of resistance gene loci from widely varied plant species suggests an important role for gene duplication and sequence divergence in the evolution of these gene families.

To understand the evolution of clusters of disease resistance genes, we have focused on the largest cluster of resistance genes in lettuce. Host–pathogen studies have identified at least 15 dominant resistance genes (*Dm* genes) in lettuce (Kesseli et al., 1994; Witsenboer et al., 1995) that match dominant avirulence genes in lettuce downy mildew (*Bremia lactucae*; Ilott et al., 1989). Most of these resistance genes map to three major clusters (Kesseli et al., 1994; Witsenboer et al., 1995). The largest cluster contains at least 10 *Dm* resistance genes as well as a gene for resistance to root aphid (Farrara et al., 1987; Bonnier et al., 1994; T. Nakahara and R.W. Michelmore, unpublished data). In lettuce cultivar Diana, this cluster includes the downy mildew resistance genes *Dm3* and *Dm1*, which are separated by 9 centimorgans (Kesseli et al., 1994). Several low-copy randomly amplified polymorphic DNA (RAPD) and restriction fragment length polymorphism (RFLP) markers near *Dm3* are components of localized duplications; long-range restriction mapping indicated that the region containing most of these duplications spanned 1.5 Mb (Anderson et al., 1996). A linear order of deletion breakpoints was determined based on the presence or absence of these duplicated markers in nine fast neutron–induced *Dm3* mutants (Okubara et al., 1994; Anderson et al., 1996). Sequences similar to the NBS of resistance genes were amplified from lettuce by use of degenerate oligonucleotides. One of these sequences mapped to the *Dm3* locus and detected a localized family of sequences. These resistance gene candidates (RGCs) from the *Dm3* locus (*RGC2*) encode an LRR region in addition to the NBS (Shen et al., 1998). Nine full-length copies of this gene, ranging in size from 7 to >13 kb, have been sequenced and analyzed in detail (Meyers et al., 1998).

Twenty-two members of the *RGC2* gene family and intergenic sequences from the *Dm3* region were characterized in this study. The region was dissected using deletion mutants, densely spaced molecular markers, and a collection of lettuce bacterial artificial chromosome (BAC) clones containing the *RGC2* gene family. Eighteen *RGC2* sequences could be positioned using the deletion breakpoint map. Sequence comparisons and phylogenetic analyses within the *RGC2* family demonstrated that these sequences exhibit a high level of sequence diversity. A size estimate for the region suggests that the duplicated sequences comprise an extensive gene family spread over >3.5 Mb. Resistance gene loci may represent one of the largest and most diverse types of gene families in plants.

RESULTS

Localization of Additional Molecular Markers on the Deletion Breakpoint Map

The original deletion breakpoint map was constructed using a panel of nine fast neutron–induced *dm3* mutants to locate 12 markers, including markers such as $AC15_{800}$, that were used in this study (Anderson et al., 1996). None of these markers was missing in all of the deletion mutants, indicating that we had not saturated the region, the markers were duplicated, or the closest breakpoints overlapped within the *Dm3* gene and no region was missing in all the mutants. Subsequently, two markers, microsatellite MSAT15-34 and the hybridization marker $IPCR_{800}$, were identified from the sequence flanking a *dm3* T-DNA insertion mutant that indicated a region was missing in all the mutants (Okubara et al., 1997). Both of these markers are duplicated within the *Dm3* region; one copy of each marker was missing in all mutants.

To identify additional markers, we saturated *Dm*-containing regions with RAPD and amplified polymorphism (AFLP) markers. Initially, 336 RAPD primers and 80 AFLP primer pairs were screened against a panel of 12 genotypes. Six templates represented pooled DNA samples for bulked segregant analysis (Michelmore et al., 1991) of each of the major *Dm* clusters plus DNA samples of six fast neutron–induced *Dm* mutants. Most of the markers identified with this panel were polymorphic between the bulked samples but were not missing in the deletion mutants; therefore, they were only loosely linked to *Dm* genes (Figure 1). Consequently, 500 more RAPD primers and 648 AFLP primer pairs were screened against a panel of DNA samples from four fast neutron–induced *Dm* mutants. The largest deletion mutant, *dm3r1608* (Okubara et al., 1997), was used to identify markers missing in the *Dm3* region. Markers missing in *dm3r1608* were then mapped on the complete panel of nine *Dm3* deletion mutants. A total of ~58,000 RAPD and AFLP loci were screened: 9000 RAPD loci, 9000 AFLP loci by using the panel of 12 DNAs, and 40,000 loci by using the panel

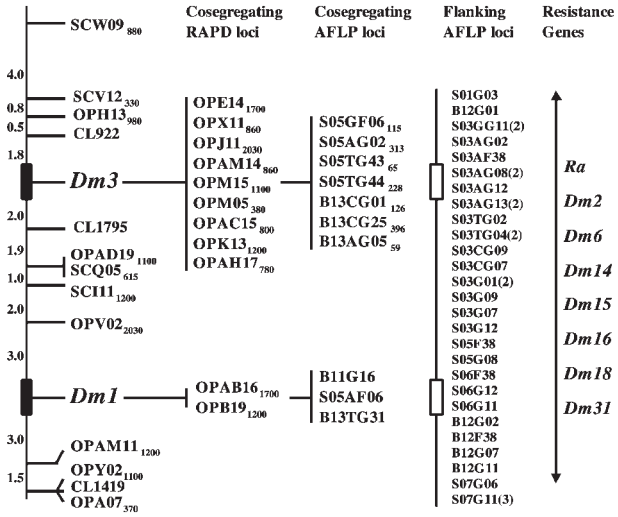

Figure 1. Genetic Map of the Major Cluster of Resistance Genes in Lettuce.

The genetic map at the left was derived from the analysis of an intraspecific cross using cultivars Calmar × Kordaat (Kesseli et al., 1994). Distances are in centimorgans. RAPD and AFLP loci were identified using bulked segregant analysis and the deletion mutants of cultivar Diana. Polymorphic fragment sizes (in base pairs) are shown in subscript. Cosegregating RAPD and AFLP loci that were missing in multiple deletion mutants are shown in the second and third columns. The precise genetic location of AFLP markers that differentiated resistant and susceptible bulked segregants for both *Dm1* and *Dm3* but were present in all deletion mutants was not determined (fourth column). Numbers within parentheses next to AFLP markers indicate multiple polymorphic bands identified using this primer pair. The resistance genes at far right have been shown by classic genetics to be linked to this cluster (Farrara et al., 1987; Maisonneuve et al., 1994; R.W. Michelmore, unpublished results). Filled boxes indicate that cosegregating markers are found adjacent to *Dm1* and *Dm3*; open boxes indicate that flanking AFLP markers are nearby but not close enough to be missing in deletions surrounding these genes.

of four mutants. This level of saturation should have provided an average spacing of one marker every ~50 kb through the genome. Seven markers from different primer pairs were identified as missing in dm3r1608 and mapped on the complete panel. One AFLP band, B13CG01, was missing in all mutants, and the remaining six markers mapped to the left of *Dm3* (Figure 2).

Genomic DNA gel blots using the NBS-encoding region from *RGC2B* identify numerous copies of this sequence that were missing in the *dm3* deletion mutants (Shen et al., 1998). Most bands were localized in the *Dm3* region in a manner consistent with our deletion map (Figure 2). One band, NBS2B band L, gave a pattern apparently inconsistent with the map. However, as previously noted by Anderson et al. (1996), multiple identical copies of a marker would obstruct correct mapping. When markers are duplicated on one side of *Dm3*, the proximal copy (relative to *Dm3*) is masked by the presence of the distal copy. When markers are duplicated on both sides of *Dm3*, a more complicated mapping pattern is observed. NBS2B band L is most parsimoniously explained by the presence of two copies flanking *Dm3* (Figure 2). A similar situation exists for $AC15_{800}$ band C, which was incorrectly mapped previously (Anderson et al., 1996); this hybridization pattern can most easily be explained by the presence of at least two copies positioned as shown in Figure 2.

Toward the end of the study, the isolation and hybridization of the region flanking *RGC2B* identified a low-copy probe from the end of a λ clone. 651END is an 800-bp fragment 4 kb 3′ to the end of the 3′ untranslated region in *RGC2B.* The utility of this sequence was assessed by using it as a probe in genomic DNA gel blot analysis; it proved to be a highly informative marker. Hybridization with 651END identified at least 14 distinct fragments from wild-type lettuce genomic DNA (Figure 3). These copies were all consistent with the deletion map.

In summary, >90 markers localized around *Dm3* have now been identified. Four of these markers were missing in all of the deletion mutants and therefore colocalize with the *Dm3* gene.

Isolation and Analysis of BAC Clones in the *Dm3* Region

Forty-eight BAC clones were identified by use of the *Dm3*-specific RFLP markers $AC15_{800}$ or the NBS-encoding region of *RGC2* (NBS2B) as hybridization probes. Secondary screens detected both markers in all selected BAC clones; in retrospect, this was not surprising, because both markers were later shown to be from within the *RGC2* gene (Figure 4A). Hybridization, sequence data, and the microsatellite MSATE6 indicated that *RGC2* genes were present as single copies in 47 of the 48 BACs. The largest BAC, BAC H1 (210 kb), contained two copies of MSATE6, $AC15_{800}$, and the NBS-encoding region of *RGC2* (data not shown); sequence data confirmed the presence of two copies of these markers. The average length of the BACs containing the *RGC2* genes was 120 kb. Therefore, the average spacing between *RGC2* genes is at least this distance.

The 48 BAC clones were initially assembled into 21 groups according to $AC15_{800}$ and NBS hybridization patterns. Groups of BACs were named according to the *RGC2* sequence contained on those BACs. A unique $AC15_{800}$ and NBS2B banding pattern defined each group. Sequence analysis and the MSATE6 bands confirmed 20 of the BAC groups, although sequence data divided one group identified by hybridization analysis in two. Two regions of the *RGC2* gene were sequenced from each BAC: a 1.4-kb region that contains the NBS and an 800-bp region comprising the $AC15_{800}$ marker (Figure 4B and Table 1). It was subsequently found that the $AC15_{800}$ marker spans an intron–exon boundary;

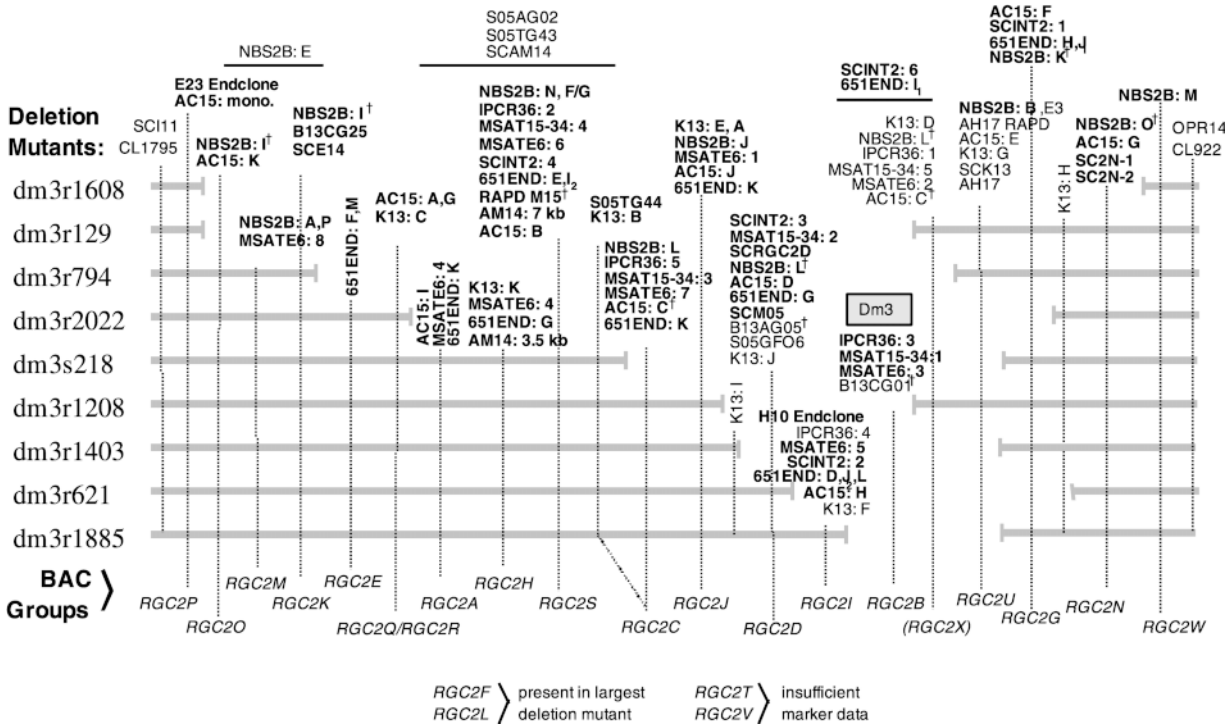

Figure 2. Deletion Breakpoint Map with Positions of *RGC2*-Containing BACs.

The mutants are ordered according to their left breakpoints. Genomic regions present in each mutant were inferred by the presence of the markers. Positions of *RGC2* BAC groups (Table 2) are given below the map. Markers in boldface type are present on the BACs indicated below by the connecting dotted line. Markers in normal typeface mapped to the region but were not detected on a BAC. The positions of markers that could not be located precisely are shown above bars at top. One $AC15_{800}$ band that was monomorphic (mono.) when analyzed by DNA gel blot hybridization (Anderson et al., 1996) was mapped by an analysis of recombinant progeny (D.B. Chin and R.W. Michelmore, unpublished data). Individual bands of multicopy markers are denoted by a colon and the letter(s) or number designating the individual band. Identical duplicate markers, which were detected by their presence on nonoverlapping BACs, are noted by a dagger. The order could not be determined when several groups of BACs mapped within the same breakpoints. Data modified and updated from Anderson et al. (1996).

the $AC15_{800}$ marker only amplified from 17 BAC groups because of variable sequences in the intron. Only one of the two *RGC2* copies could be amplified and sequenced from BAC H1 (*RGC2Q* but not *RGC2R*). Therefore, the 48 clones contained a total of 23 different copies of the *RGC2* gene (Table 2).

The 48 BAC clones represented nearly all of the copies of *RGC2* present in the genome. An average of ~2.2 BAC clones were identified for each *RGC2* copy; this was consistent with the two to three times genomic coverage calculated for the libraries (Frijters et al., 1997; Z. Zhang and R.W. Michelmore, unpublished data). Hybridizations with $AC15_{800}$ and the NBS-encoding region of *RGC2B* identify at least 10 and 18 bands, respectively (Anderson et al., 1996; Shen et al., 1998). Several lines of evidence indicated that one *RGC2* copy (*RGC2X*) had not been cloned: four groups of BACs containing $IPCR_{800}$ and the associated microsatellite MSAT15-34 were identified rather than the five expected; one copy of each of the markers NBS2B, $AC15_{800}$, SCINT2, and MSATE6 was identified in genomic DNA but was not present on a BAC clone. All of these missing copies mapped adjacent to *Dm3* in the deletion mutants (Figure 2). The colocalization of the markers absent from the BAC clones indicates that there are few, if any, additional *RGC2* members that were not cloned. Therefore, there are probably 24 copies of the *RGC2* gene in cultivar Diana.

Positioning of BAC Groups in the *Dm3* Region by Using the Deletion Mutant Map

Two approaches were used to position the BAC groups. Molecular markers from the *Dm3* region were assayed on the BAC clones, and markers derived from *RGC2* sequences on the BACs were located on the deletion breakpoint map. The use of multiple markers to localize the BAC groups reduced the likelihood of misplacement due to duplications. Two multicopy polymerase chain reaction (PCR)–

based markers, MSATE6 and SCINT2, were developed from sequence comparisons between *RGC2* genes. Of the 48 clones, only two BACs (containing *RGC2T* and *RGC2V*) could not be positioned unequivocally because of insufficient marker information; these BACs did not contain markers that map immediately adjacent to *Dm3* and have not been completely characterized. Sixteen of the 22 BAC groups, including BAC H1, which contains two *RGC2* genes, were positioned between the breakpoints in the largest deletion mutant (Figure 2). In addition, four groups of BACs contained *RGC2* markers that were outside the region missing in our largest deletion mutant; of these, two have been mapped to the left or right of *Dm3* by use of recombinants selected from a large F_2 population segregating for *Dm3* (Figure 2; D.B. Chin, unpublished data).

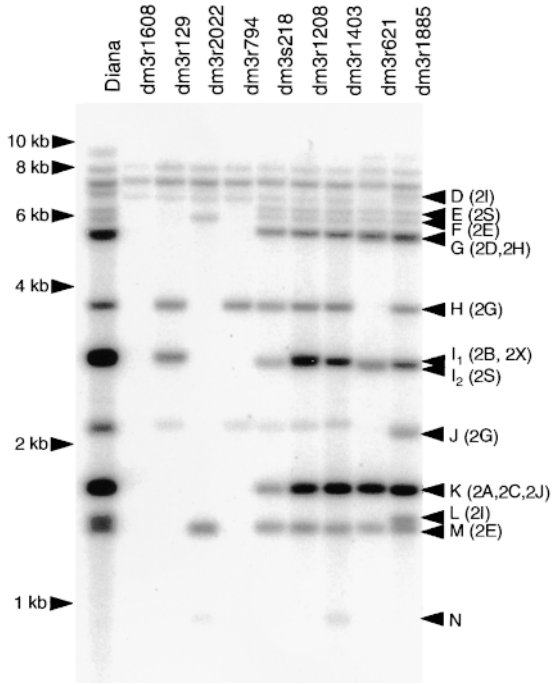

Figure 3. Hybridization of the Region 3′ to *RGC2B* to the Panel of Deletion Mutants.

Autoradiography showed variable loss of markers in deletion mutants that correspond with particular members of the *RGC2B* gene family. Genomic DNA from wild-type Diana and the nine deletion mutants was cut with HindIII and probed with 651END, which is an 800-bp fragment ~4 kb 3′ to the poly(A) site of *RGC2B*. The arrowheads to the right indicate mapped loci (Figure 2), designated by letters; the *RGC2* family member(s) that corresponds to each band is noted within parentheses with a 2. The three high molecular weight bands were not identified among the BAC clones. Band N was observed in dm3r2022 and dm3r1403 but not in wild-type Diana and was not identified among the BAC clones. Positions of DNA standards are indicated at left in kilobases.

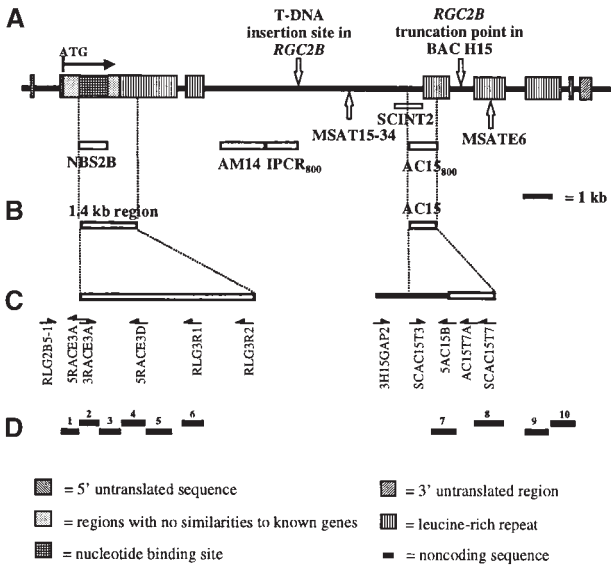

Figure 4. Genomic Structure of *RGC2* Homologs Showing Positions of Relevant Markers.

(A) Diagram of a typical *RGC2* gene. Markers AM14 and $AC15_{800}$ (Anderson et al., 1996), $IPCR_{800}$ and MSAT15-34 (Okubara et al., 1997), and MSATE6, SCINT2, and NBS2B were used to map the deletion breakpoints and BAC groups (Figure 2).

(B) The 1.4-kb region in the 5′ end and $AC15_{800}$ in the middle of the coding region were sequenced for phylogenetic analysis and to confirm the identity of the sequences contained on the BACs.

(C) Relative positions of the primers used for sequencing and for amplifying markers (Table 1). Half arrows indicate positions and directions of primers.

(D) The gene segments (10) used for phylogenetic analysis to examine the nine *RGC2* gene sequences for evidence of recombination and gene conversion.

Estimation of the Size of the *Dm3* Region

We estimated the degree of overlap between BACs within a group by using HindIII digests and AFLP fingerprints. AFLP fingerprinting using primers with one discriminatory base produced an average of 10 fragments per primer pair per BAC (Meyers, 1998). AFLP fingerprints frequently showed identical bands present in diverse BACs that mapped to nonadjacent locations consistent with a high level of duplication throughout the region. AFLP fingerprints confirmed the BAC groupings generated with other markers; however, they did not allow the unambiguous identification of overlaps between groups because it was impossible to determine the difference between duplications and overlaps.

HindIII digests of BAC clones provided a more comprehensive analysis of each group of BACs. The size of each group of BACs was determined by summing the unique and common fragments within a group (Table 2). Groups determined to be potential neighbors on the basis of their position on

Table 1. Oligonucleotide Primers Used for PCR[a]

Purpose and Origin	Designation	Sequence (5′ to 3′)
Amplification and sequencing of 1.4-kb region	3RACE3A	GCCTTGTGTGGGATGGGTGGA
	RLG3F1	GAAACGAGCTACCACAATCTCC
	RLG3R1	GAAACCTTAGCGACTTATCTCCA
	RLG3R2	GAACGCTCTGCCATCTCATTG
	5RACE3D	GCTCGTTTCAAAGACTTTGGGC
Amplification and sequencing of $AC15_{800}$ fragment	SCAC15T7	CCGTGAGAGGTGAAAACAAGTA
	SCAC15T3	GTGAGACCGTGACTTGGATG
	AC15T7A	GGTGTGAGGTTGTGGAATGG
Isolation of BAC end clones	ENDCLN1	CCTAAATAGCTTGGCGTAATCATG
	ENDCLN2	TGACACTATAGAATACTCAAGCTT
	ENDCLN3	CGACCTGCAGGCATGCAAGCTT
	ENDCLN4	ACTCTAGAGGATCCCCGGGTAC
	ENDCLN5	TTTTACAACGTCGTGACTGGG
Amplification of marker MSATE6	5MSATE6-1	CCCAAGAAGAATCCTACCA
	3EXON4C	AGTGATTGTGAAGAAGGAAGAA
Amplification of marker SCINT2	3H15GAP2	GTTTGGGCTGAACGGGA
	5AC15B	ATGGAAAATGCCAACAGCT
Amplification of marker SCRGC2D	3E14GAP1	CAAATCATTCCAGGCTCTCA
	5AC15B	ATGGAAAATGCCAACAGCT
Amplification of marker SCRGC2B	RLG2B5-1	GATCAGAAGAGACTGTTCACAC
	5RACE3A	CACACAAGGCTACCATGTGGA

[a] Positions of primers within *RGC2* genes are shown in Figure 4C.

the deletion breakpoint map were examined for overlapping fragments. However, there was not sufficient overlap or resolution to identify unambiguous small overlaps between groups. Hybridization with either whole BACs or end clones did not provide additional resolution or further evidence of sequence duplication throughout the region. Although small overlaps cannot be excluded, large overlaps between BAC groups were not present.

The size of the region containing the duplicated *RGC2* sequences was estimated by summing the sizes of the BAC groups. Assuming little or no overlap between 22 BAC groups, the region that we have cloned must be at least 3 Mb (Table 2). Twelve of these BAC groups were ⩾150 kb; however, only one BAC contained two copies of *RGC2.* This was the largest *RGC2*-containing BAC (H1; 210 kb), suggesting a minimum spacing of 150 kb between *RGC2* copies. The BAC group containing *RGC2I* and *RGC2B* (BACs H10 and H15; see below) determined that these copies are between 100 and 255 kb apart. Both lines of evidence indicate that the average distance between *RGC2* copies is at least 145 kb. Therefore, the 24 members of the *RGC2* gene family span a minimum of 3.5 Mb. This is significantly greater than the previous estimate of 1.5 Mb based on summation of long-range restriction fragments detected by $AC15_{800}$ (Anderson et al., 1996). The difference may result from the lack of detection of all *RGC2* copies in the earlier analysis; approximately half of the fragments detected by $AC15_{800}$ in our subsequent analysis of BAC clones were not present in conventional genomic DNA gel blots. The estimate of 3.5 Mb is conservative; the region containing the *RGC2* multigene family may be considerably larger because of gaps between the BAC groups.

Identification of *Dm3* Candidate Sequences

One of the goals of this project was the identification of candidate sequences for the *Dm3* resistance gene. Several lines of evidence indicate that *RGC2B* is *Dm3.* Markers specific to *RGC2B* were missing in all deletion mutants. Several of these markers were present in the single BAC (H15) containing *RGC2B.* The marker SCRGC2B (Table 1) is amplified specifically from the 5′ sequence of *RGC2B* and is missing in all deletion mutants. The insertion of a T-DNA element into *RGC2B* correlated with a loss of *Dm3* function (Okubara et al., 1997). Although SCRGC2B, the T-DNA insertion site, and other *RGC2B*-specific markers were present on BAC H15, sequence analysis indicated that this BAC did not contain the complete gene (Figure 4A; Meyers et al., 1998). Therefore, a genomic λ library was screened, and a 20-kb clone was isolated that contained the complete *RGC2B* gene as well as ~4 kb of both upstream and downstream sequences. Transgenic complementation is currently under way with this clone.

The region missing in all deletion mutants was characterized in detail to determine whether there were additional *RGC2* sequences that could be candidates for *Dm3.* Ends of the BACs in the region immediately adjacent to *Dm3* were isolated by using inverse PCR. Overlap was detected between BAC H10 (carrying *RGC2I*) and BAC H15 (carrying *RGC2B*) by amplification and DNA gel blot hybridization with these end clones. One end clone from BAC H10 was a single copy in the region, present in BAC H15, and present in deletion mutant dm3r1885 (Figure 2). One BAC H15 end clone was duplicated on several BACs; the corresponding fragment from BAC H10 was sequenced to distinguish between identity and duplication of closely related sequences. The BAC H10 copy was identical to that of BAC H15, whereas copies from other BAC clones were not. Therefore, BAC H10 and BAC H15 represent contiguous sequence between *RGC2I* and *RGC2B.* All *RGC2I*-specific markers were present in deletion mutant dm3r1885 (Figure 2). Therefore, *RGC2I* is not *Dm3.* As discussed above, one *RGC2* gene that was not present in our BAC library, *RGC2X*, is adjacent to *RGC2B.* Therefore, the gap between *RGC2B* and the right-hand breakpoints in mutants dm3r129 and dm3r1208 was not present on a BAC clone and could not be analyzed. However, markers specific to *RGC2X* are present in both dm3r129 and dm3r1208 (Figure 2); therefore, *RGC2X* is unlikely to be *Dm3.*

Determination of Gene Density in the *Dm3* Region

Random and low-copy DNA fragments from the *Dm3* region were sequenced to search for other genes duplicated locally in the *Dm3* region. A total of 15.7 kb of sequences resulted from the sequencing of end clones from BACs H1, H5, H10, H15, H149, E7, E29, and E33. In addition, low-copy fragments were identified by reverse genomic DNA gel blot hybridizations to libraries made from partial digests of three BACs (H15 [*RGC2B*], E32 [*RGC2H*], and H2 [*RGC2A*]) by using total genomic DNA as a probe. These BACs represented a total of 425 kb of genomic DNA, including ~90 kb from BAC H15 upstream of the *Dm3* candidate, *RGC2B*; BACs H2 and E32 contained diverse *RGC2* sequences that mapped to the left of *Dm3* (Figure 2). DNA sequences totaling ~50 kb were obtained from 93 sequencing reads of clones containing low-copy DNA from the three BAC subclone libraries. The frequency of reads with similarity to the *RGC2* gene on each BAC served as a check for saturation of sampling for low-copy sequences. Twenty-nine low-copy subclones were identical to the sequence in or 5′ of the full-length *RGC2* genes (sequence described in Meyers et al., 1998). In total, these sequences represented 47% of the complete *RGC2* gene and the 5′ flanking region. Therefore, any additional duplicated and sizeable genes should have been detectable with this level of sampling.

Table 2. BAC Groups Constructed on the Basis of Shared *RGC2* Family Members and Other Markers

RGC2 Designation	GenBank Accession No.	BACs in Group	Lengths of BACs (kb)	Length of BAC Group (kb)
RGC2A	AF072268	H2, H8, H148	125, 70, 150	160
RGC2B	AF072267	H15	100	100
RGC2C	AF072269	H18, H19, E42, H200	50, 75, 120, 90	135
RGC2D	AF072270	E14	140	140
RGC2E	AF072276	H9, H12, H98, E15, E38, E58	125, 150, 50, 100, 100, 125	>200
RGC2F	AF072277	E12, H203, H147	170, 90, 45	170
RGC2G	AF072278	E61, E37, E46	150, 100, 120	150
RGC2H	AF072279	H3, E55, E32, E139	140, 170, 110, 140	185
RGC2I	AF072280	H10	160	160
RGC2J	AF072271	E6, E49	130, 130	130
RGC2K	AF072272	E22, E51, E52	140, 150, 140	155
RGC2L	AF072281	E11	160	160
RGC2M	AF072282	H5	50	50
RGC2N	AF072273	E29, E33, H210	100, 150, ND[a]	>200
RGC2O	AF072274	E7, E36, E54, H208	90, 60, 150, 50	175
RGC2P	AF072283	E23	150	150
RGC2Q	AF072284	H1, H4	210, 175	210
(RGC2R)	(Not sequenced)	(H1)[b]	(210)[b]	(210)[b]
RGC2S	AF072275	H201	100	100
RGC2T	AF072285	H209	45	45
RGC2U	AF072286	H149	140	140
RGC2V	AF072287	H207	ND	ND
RGC2W	AF072288	E155	140	140
RGC2X	NA[c]	Not present in BAC library	See text	NA

[a] ND, not determined.
[b] Second copy present on BAC H1.
[c] NA, not available.

The only significant homologies found were to mobile elements and plant cyclin genes. Retrotransposable elements were identified by four subclones that had significant similarity to the copia-like class of elements recently isolated from intergenic regions of the maize genome (best BLASTX score = 122; $P = 2 \times 10^{-30}$; SanMiguel et al., 1996). One clone identified transposable element sequences in the databases, with BLASTX scores of 79 to the *TNP2* element of snapdragon (Antirrhinum; GenBank accession number X57297) and 69 to the maize transposon *En-1* (GenBank accession number S29329). Retrotransposon-like sequences (best BLASTX score = 144; $P = 8 \times 10^{-46}$ to GenBank accession number 226407) also were identified in three end clones: one end of each of the BACs H1, H5, and H15. BLAST searches using the remaining end clone sequences found no significant similarity to sequences in the databases. Hybridization using the retroelements probed to the BACs in the *RGC2* groups indicated that retroelement copy number varied from low to high copy in the *Dm3* region (data not shown). Sequences with similarity to plant cyclins, proteins involved in cell cycle regulation, were identified from subclones of two BAC sequences (best BLASTX score = 68.5; $P = 2 \times 10^{-21}$ to GenBank accession number X82036). These clones contained four regions spanning 750 bp that are conserved in cyclins; however, the conserved sequences were interspersed with stop codons, and one clone contained a poly(A) motif, suggesting that this sequence could be an ancient processed pseudogene.

Rapid amplification of cDNA ends (Frohman et al., 1988) analysis of lettuce cDNA by using primers from the cyclin sequences detected no transcripts (data not shown). Hybridization to the 48 *RGC2*-containing BACs indicated that the cyclin fragment was present in only three copies in the *Dm3* region (data not shown). Therefore, the cyclin-related sequences in the *Dm3* region are unlikely to be functional. No open reading frames >300 bp were found in the majority of low-copy sequences, and there was no significant similarity to sequences in the databases. Consequently, there was no evidence for additional functional genes on these BACs; however, a single-copy or small gene could have been missed in this analysis.

Sequence Comparisons of the *RGC2* Family

Segments from the majority of the *RGC2* copies were sequenced to determine the evolutionary relationship within the multigene family and to investigate the genetic mechanisms shaping the cluster. Initially, we sequenced two segments: an 800-bp 3′ region corresponding to the $AC15_{800}$ marker and a more 5′ 1.4-kb region encoding the NBS and the first six LRRs (Figure 4C). Eighteen unique $AC15_{800}$ sequences were obtained first and used to group the BACs based on sequence identity (see above). $AC15_{800}$ did not amplify from the remaining four groups of BACs. The $AC15_{800}$ marker was later shown to span the intron 3/exon 4 boundary of the *RGC2* gene, including ~400 bp of the LRR-encoding region in exon 4 (Figure 4C). Twenty-two sequences composed entirely of open reading frame were obtained for the segment containing the NBS homology. This segment was used for pairwise comparisons and phylogenetic analysis of the *RGC2* family.

Pairwise comparisons of the 1.4-kb sequences demonstrated a high level of diversity within the *RGC2* gene family. The nucleotide sequence identity ranged from 53 to 97% between the 22 copies (Table 3). The phylogenetic relationships of the *RGC2* family members were determined by use of several tree-building methods, all of which gave nearly identical results; neighbor-joining (Figure 5A) and maximum parsimony (Figure 5B) trees are shown. Bootstrap values indicated that the *RGC2* family was composed of several well-supported subfamilies, single divergent members, and some less well supported subfamilies (Figure 5B). The subfamily containing *RGC2B* (the *Dm3* candidate), *RGC2S*, *RGC2C*, and *RGC2D* (as well as the missing copy *RGC2X*) was also supported by other marker data, notably the presence of $IPCR_{800}$ and MSAT15-34.

There was no correlation between the phylogenetic relationships and physical position of the *RGC2* sequences (Figure 5B). Closely related sequences mapped to different, nonadjacent locations on the deletion breakpoint map (Figure 2). The complexity of the relationship between phylogeny and physical position indicated that many rearrangements have occurred during the evolution of this region. Therefore, it is impossible to reconstruct the history of duplications and deletions in the *Dm3* region from the analysis of a single haplotype.

Trees generated from different regions of the *RGC2* gene were then compared for evidence for chimeric genes that could have resulted from gene conversion or unequal crossing over between family members. Initially, phylogenetic trees were generated using the $AC15_{800}$ segment from 18 *RGC2* copies (data not shown); however, the bootstrap values were low, and therefore, the trees were not sufficiently robust to detect unequal crossing over. Toward the end of the study, we obtained the full-length sequences for nine *RGC2* genes (Meyers et al., 1998). The full-length sequences were aligned and divided into ~600-bp segments, taking into account functional domains and the position of introns (Figure 4D). Phylogenetic trees were generated for each segment and compared. Trees generated for segments 5′ to the large intron (segments 1 to 6, Figure 4D) were robust trees, with bootstrap values for the majority of nodes >70% and often >90% (Figure 6).

The phylogenetic relationship for most genes was invariant, indicating that there has been no unequal crossing over or significant lengths of gene conversion between these genes. However, the phylogenetic relationship of *RGC2S* and *RGC2B* relative to *RGC2C* and *RGC2D* or *RGC2J* changed dramatically between the first and fourth segments (Figures 6A and 6B). Inspection of the sequences for these genes indicated that a crossover had occurred between

Table 3. Nucleotide Sequence Identity and Amino Acid Similarity of a 1.4-kb Region of the *RGC2* Sequences[a]

	RGC Sequence																				
	2A	*2B*	*2C*	*2D*	*2E*	*2F*	*2G*	*2H*	*2I*	*2J*	*2K*	*2L*	*2M*	*2N*	*2O*	*2P*	*2Q*	*2S*	*2T*	*2U*	*2V*	*2W*
RGC2A		69.4	65.2	66.7	66.2	69.6	66.2	65.8	64.5	63.9	63.3	66.6	66.6	95.1	67.5	68.7	65.4	71.4	64.5	66.2	62.8	65.2
RGC2B	75.7		81.6	84.8	76.0	76.3	63.3	64.4	62.9	62.2	65.6	78.2	69.4	68.4	77.0	67.8	68.1	94.0	63.8	66.2	61.2	67.1
RGC2C	72.7	86.6		95.1	76.4	73.9	61.9	62.1	61.0	60.1	62.5	74.5	69.4	64.8	75.8	66.7	65.3	79.1	60.4	61.4	59.1	61.6
RGC2D	72.2	87.1	95.5		76.2	75.7	63.0	62.9	61.7	60.9	64.9	74.1	68.6	66.2	77.0	66.7	65.5	81.7	63.2	61.7	59.6	62.5
RGC2E	74.3	82.5	81.9	81.7		78.5	62.3	62.3	61.5	60.9	62.3	80.2	67.4	66.6	83.0	68.0	67.0	78.5	60.2	61.7	59.6	61.6
RGC2F	74.8	81.6	78.9	79.1	83.0		63.8	64.4	62.8	63.3	64.7	75.2	68.8	68.5	76.1	67.8	68.1	79.7	62.4	64.4	62.4	63.3
RGC2G	72.5	69.8	69.2	68.6	66.7	67.9		96.6	94.3	94.1	59.6	62.0	56.9	64.6	62.8	65.5	59.8	64.4	59.6	58.6	92.8	57.0
RGC2H	72.6	71.5	69.2	69.6	66.4	67.7	97.3		93.6	92.4	60.4	63.3	56.9	65.0	62.9	65.6	61.2	65.0	59.7	56.3	93.0	56.2
RGC2I	71.6	70.9	68.7	69.5	66.6	67.3	95.1	96.1		91.5	58.7	62.2	55.2	63.6	62.4	64.2	59.4	63.5	58.4	55.7	90.3	55.6
RGC2J	68.0	66.1	65.2	66.0	64.2	65.4	92.0	92.7	90.9		59.7	61.3	55.5	63.5	60.9	65.3	59.2	63.5	58.8	54.8	89.0	53.7
RGC2K	73.5	71.6	70.6	71.0	71.4	71.4	63.3	63.5	63.0	61.0		62.9	56.1	62.3	62.8	64.9	59.8	66.3	60.3	56.3	57.4	56.4
RGC2L	73.2	85.3	80.7	81.1	85.2	80.0	65.5	66.0	66.1	64.2	69.5		63.5	66.2	81.5	65.8	67.9	81.1	60.7	60.5	59.1	59.5
RGC2M	77.8	78.0	75.3	72.8	77.0	77.9	61.2	61.5	61.7	59.1	69.6	72.5		65.4	69.1	60.1	62.9	70.3	56.7	61.4	53.5	60.6
RGC2N	96.2	74.7	72.8	72.2	74.6	74.6	72.1	73.0	72.0	67.9	72.8	73.2	76.9		67.4	67.8	64.4	70.6	64.5	66.5	61.8	65.2
RGC2O	74.9	80.2	82.2	81.7	87.5	80.7	68.9	70.5	69.1	65.9	72.2	85.5	77.1	74.5		69.4	67.4	80.8	63.8	63.8	59.4	62.5
RGC2P	74.1	72.6	69.9	69.4	72.4	72.2	71.5	70.5	69.8	69.3	71.8	73.2	69.7	73.5	75.1		64.2	69.4	64.0	58.7	62.6	57.3
RGC2Q	71.9	75.5	72.6	72.7	74.7	73.4	66.5	67.4	66.6	63.8	68.6	73.6	70.2	71.8	73.4	69.8		70.1	58.6	62.6	58.4	59.5
RGC2S	76.5	96.6	84.5	84.6	83.9	83.4	67.9	67.8	67.9	65.8	72.0	85.1	78.3	76.4	84.4	72.6	76.7		64.0	66.5	62.0	66.6
RGC2T	72.5	69.9	68.0	68.4	69.2	69.2	61.3	63.5	62.1	57.3	69.7	67.2	66.6	72.1	67.8	70.8	65.5	69.1		59.6	57.1	57.3
RGC2U	71.0	72.4	66.4	64.9	66.6	68.5	61.9	61.8	61.4	57.3	65.6	63.5	67.1	71.5	68.8	65.1	66.0	71.0	66.8		54.5	82.6
RGC2V	72.6	70.9	68.7	69.0	65.9	67.7	95.1	97.5	94.7	90.4	63.2	64.7	61.3	72.0	68.7	69.7	67.1	68.1	62.3	63.0		53.2
RGC2W	71.3	73.7	69.5	68.2	70.4	70.4	65.0	64.4	64.1	60.5	68.3	65.3	69.9	71.6	70.1	66.5	65.4	71.6	67.6	84.4	65.1	

[a] Nucleotide comparisons are shown above the diagonal markers, and comparisons of the predicted amino acid sequences are shown below the diagonal markers.

progenitor homologs somewhere within the first part of segment 2 that contains conserved domains of the NBS. This resulted in a chimeric progenitor for *RGC2S* and *RGC2B* containing 5′ sequences similar to *RGC2J* and 3′ sequences similar to the progenitor of *RGC2C* and *RGC2D.* Trees for the last four segments comprising the 3′ LRR were not robust (i.e., had low bootstrap values). Therefore, it was not possible to discern evidence of chimeric genes. The less robust trees may be the consequence of divergent selection acting on the 3′ LRR region (Meyers et al., 1998), increasing diversity between related sequences and obscuring phylogenetic relationships.

DISCUSSION

Resistance Gene Clusters Are Predominantly Composed of Arrays of Resistance Gene Homologs

The complex multigene family at the *Dm3* region is the largest resistance gene locus thus far characterized at the molecular level. Based on the size of our BAC clones, we have estimated that the ~24 members of the family span at least 3.5 Mb, with an average spacing of ~145 kb. Clusters of resistance genes from other species contain fewer members and are spaced much closer together (7 to 70 kb apart). The *M* locus for rust resistance in flax has ~15 members localized in <1 Mb (Anderson et al., 1997). The rice *Xa21* cluster has most of the eight resistance gene homologs in a locus spanning ~230 kb (Ronald et al., 1992; Song et al., 1997). In the tomato *Pto* cluster, five *Pto* homologs are spread over 60 kb (D.T. Lavelle and R.W. Michelmore, unpublished data), whereas the five members in both the *Cf-9* and *Cf-4* clusters are found within 36 kb (Parniske et al., 1997). The *I2* locus of tomato is composed of seven homologs spanning ~90 kb (Simons et al., 1998). In soybean, RGC sequences clustered near the *Rps2* and *Rmd* resistance loci are spaced approximately every 20 kb (Kanazin et al., 1996). Nine *RPP5*-related sequences in Arabidopsis are spread over 70 kb (Bevan et al., 1998). The increased number and spacing of RGCs in the *Dm3* region could be a reflection of the larger genome size of lettuce, 2.3×10^9 bp (Arumuganathan and Earle, 1991; Michealson et al., 1991; D. Galbraith and R.W. Michelmore, unpublished data) compared with these other species. It remains to be determined whether the organization observed for the *Dm3* region is typical for species with moderately sized genomes and whether species with larger genomes have correspondingly larger clusters.

Differences in genome size and intergenic spacing among plants have been attributed to an accumulation of various

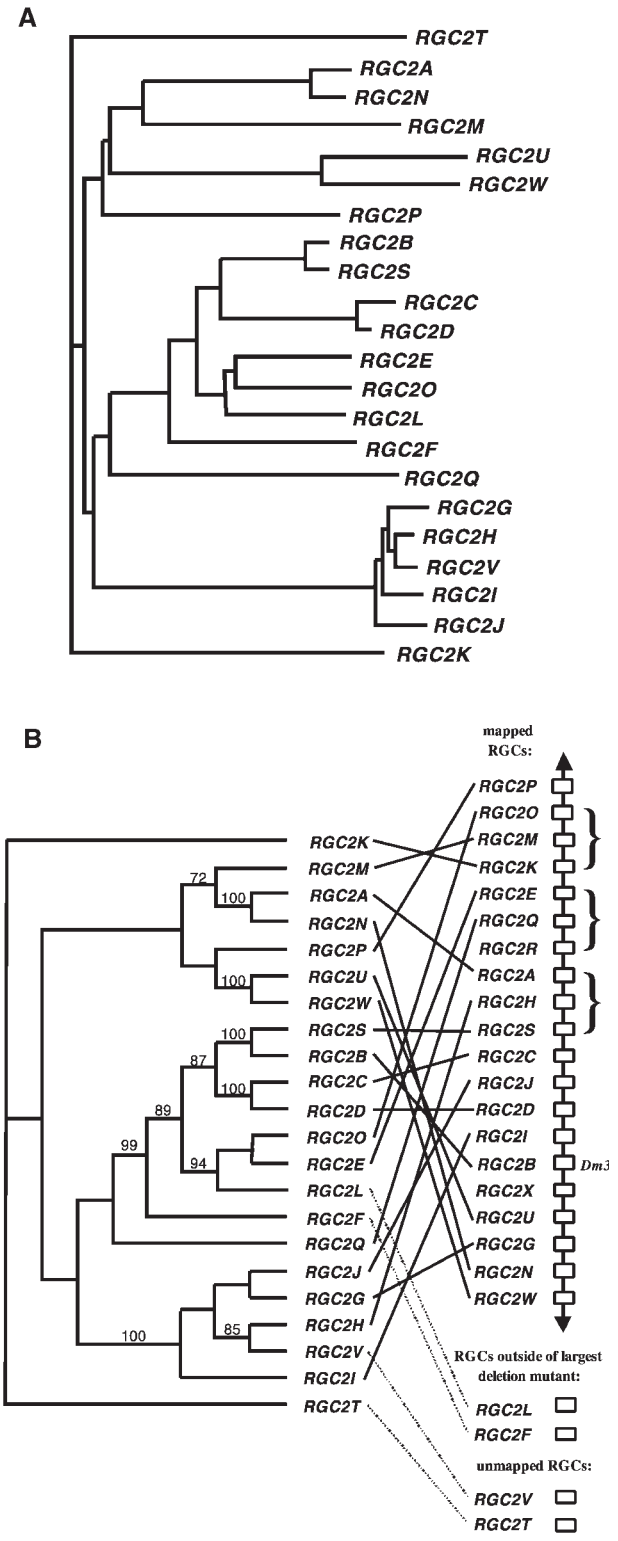

Figure 5. Phylogenetic and Physical Relationships of *RGC2* Copies.

(A) Neighbor-joining tree from distance matrices constructed according to Kimura's two-parameter method by using the DNA sequence of the 1.4-kb region that includes the NBS from each *RGC2* gene. The branch lengths are proportional to genetic distance.

(B) Maximum parsimony tree using the same DNA sequences as given in **(A)**. Bootstrap values are indicated as calculated for nodes supported with >70% of 100 replicates. Branch lengths are not scaled. The vertical line to the right represents the chromosome with the mapped positions of the *RGC2* copies in the *Dm3* region. A bracket next to the *RGC2* copies indicates uncertainty in gene order.

types of transposable elements in plant species with larger genomes (Wessler et al., 1995; SanMiguel et al., 1996). Mobile genetic elements, particularly retrotransposons, may comprise a significant portion of most plant genomes. In rice, transposon-like elements both flank and interrupt some members of the *Xa21* gene family (Song et al., 1997). The *RGC2* family in lettuce is interspersed with sequences similar to a variety of mobile elements, although there was no evidence for these sequences within *RGC2* genes (this study; Meyers et al., 1998). Transposons could have played a role in the evolution of the *Dm3* region through expansion or translocation of RGC sequences. Mobile elements can cause either localized duplications and deletions of genes (Walker et al., 1995) or more extensive chromosomal rearrangements (Fedoroff, 1989). Variation in *RGC2* copy number within the *Dm3* region occurs in diverse lettuce germplasm, detected by hybridization with $AC15_{800}$ (Anderson et al., 1996). Furthermore, sequences related to triose phosphate isomerase are linked to at least three separate *Dm* clusters (Paran and Michelmore, 1993), suggesting that these clusters resulted from duplications of an ancient cluster. Determination of the role of transposons in the evolution of resistance gene clusters requires the detailed characterization of multiple haplotypes.

There was no evidence for functional genes in the *Dm3* region other than *RGC2* homologs and transposon-related sequences. Limited localized duplications may have resulted in the cyclin-related fragments that were observed in several BACs. The *Cf-4*/*9* homologs in tomato are interspersed with fragments of *Lox* genes, which may have played a role in the duplication of that region (Parniske et al., 1997). The *Pto* cluster of protein kinase homologs also contains a single NBS-LRR gene, *Prf*, that is necessary for the function of two members of the *Pto* cluster (Salmeron et al., 1996). Although we found no evidence for duplicated large genes in the region, a single small gene, such as a protein kinase, could have been missed.

Additional *RGC2* Copies Could Encode Other Resistance Genes

The function of the ~24 *RGC2* sequences in cultivar Diana has yet to be determined. Mutation analysis indicates that

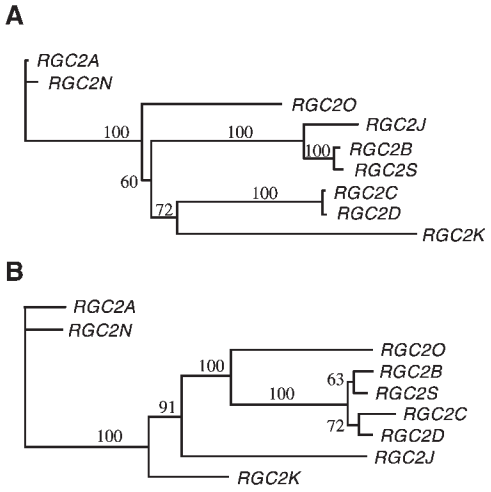

Figure 6. Phylogenetic Relationship between Segments 1 and 4 of the *RGC2* Gene.

Trees were constructed from distance matrices according to Kimura's two-parameter method by using the 10 segments of the nucleotide sequences of nine full-length *RGC2* genes (Figure 4D). Bootstrap values are indicated for each node as a percentage of 500 replicates. Comparisons among the trees demonstrated distinct differences in the trees occurring between segments 1 and 4.

(A) Phylogenetic relationship of *RGC2* genes based on the nucleotide sequence of segment 1.

(B) Phylogenetic relationship of *RGC2* genes based on the nucleotide sequence of segment 4.

only a single gene encodes *Dm3* specificity (Okubara et al., 1997). Multiple members of this family are expressed (Meyers et al., 1998) and may represent functional resistance genes. It is difficult to demonstrate the function of these genes in the absence of known pathogens detected by these genes. This is an increasing problem with the many resistance gene homologs of unknown function identified by use of PCR with primers to conserved domains as well as random expressed sequence tag and genomic sequencing (Kanazin et al., 1996; Leister et al., 1996; Yu et al., 1996; Botella et al., 1997; Aarts et al., 1998). Antisense inhibition can demonstrate the involvement of a family of sequences in resistance (Ori et al., 1997); however, it does not allow determination of individual gene activities. One possibility is to generate gain-of-function mutants to demonstrate the downstream consequence of constitutive activity in the absence of pathogen-derived ligand, as shown for *Pto* (J. Rathjen and R.W. Michelmore, unpublished data).

It is currently unknown how many of the resistance genes that map to the major cluster (Figure 1) are encoded by *RGC2* genes. Several lines of evidence indicate that the *RGC2* family comprises a tight cluster of genes. All of the *RGC2* genes that segregated in our basic mapping population (Kesseli et al., 1994) completely cosegregated with *Dm3* (Shen et al., 1998). Positioning of the BAC clones on the deletion breakpoint map indicated that *RGC2* sequences are physically clustered. Also, sequences detected by fluorescent in situ hybridization (FISH) using BAC H15, which contains *RGC2B*, are localized to a telomeric position on one chromosome (Shen et al., 1998). At least 10 distinct resistance specificities from a variety of other lettuce genotypes map to the *Dm3* locus; however, these are distinct loci, and varying amounts of recombination have been detected between *Dm3* and these other genes (Farrara et al., 1987; Kesseli et al., 1994). Even though recombinants were detected between these *Dm* genes, *RGC2* family members could encode some of these genes because the frequency of recombination in the region depends on the degree of similarity between the parents. However, it is unlikely that all are encoded by a *RGC2* family member; for example, *Dm1* (also present in cultivar Diana) maps 10 centimorgans away from *Dm3* (Kesseli et al., 1994; Figure 1). Dissimilar families of RGCs may encode the genetically distinct resistance genes linked to *Dm3.* Multiple divergent NBS-containing genes are closely linked in soybean (Kanazin et al., 1996). The evolutionary or functional significance of closely linked families of distinct NBS-LRR genes is presently unknown.

Analysis of Resistance Gene Clusters Requires the Integration of Genetic and Physical Approaches

The analysis of the complex *RGC2* region required both genetic and physical analyses. Critical components included the deletion mutant map, many molecular markers in the *Dm3* region, genomic clones containing the *RGC2* multigene family, and sequence analysis of numerous *RGC2* copies. Screens using the largest deletion mutant, *dm3r1608*, enabled us to rapidly identify additional markers in the region. The multiple deletion mutants allowed mapping of the markers and clones within a highly duplicated region in which recombination is repressed (Anderson et al., 1996) and simultaneously allowed us to refine the position of the deletion breakpoints.

Multiple markers were necessary to confirm map positions of the clones. Different types of markers varied in their usefulness for mapping the BAC clones on the deletion mutants because of differences in reliability and duplications in the *Dm3* region. Low-copy hybridization probes and the microsatellite MSATE6 were the most informative because of their specificity. These were either derived from cloned RAPD fragments or from the sequence of the full-length *RGC2* genes (Meyers et al., 1998). Random PCR-based AFLP and RAPD markers were the least reliable. Some BACs contained AFLP or RAPD markers that were duplicated elsewhere in the region. These markers sometimes amplified from a BAC clone but failed to amplify from deletion mutants predicted to contain the BAC based on more robust markers. RAPD markers were particularly informative when cloned and used as RFLPs or converted to sequence characterized amplified regions (SCARs). The duplicated

nature of the region was actually an advantage; single cloned RAPD fragments frequently provided informative markers for multiple BACs. AFLP markers were not as useful as RAPDs because they were difficult to convert to SCAR or RFLP markers. Sequencing of multiple *RGC2* genes allowed markers to be designed that detected individual BACs, specific subgroups, or all BACs.

Saturation of the region with markers as well as the development of *RGC2*-specific markers reduced the need for a contiguous set of BAC clones across the entire *Dm3* region. The physical structure of the region could be ascertained by the positional and phylogenetic analysis of the BAC groups. Marker analysis indicated that most *RGC2* copies from the *Dm3* region were contained within the BAC clones. Sequence data and a variety of informative markers allowed us to construct robust groups of BACs and to distinguish between identical and duplicated *RGC2* sequences. However, fingerprinting the BACs with AFLPs and HindIII fragments was not sufficiently informative to identify unambiguously small overlaps between groups of BACs. To identify genuine overlaps requires the development of informative PCR-based markers for the ends of each group of BACs, as was done for BACs H10 and H15.

A large array of resistance genes provides a variety of opportunities for plants to counter the challenge of variable pathogens. Pathogen populations are under selection to evade detection by a resistance gene and gain access to the host plant. There are two ways that arrays of resistance genes could be advantageous. Clusters could either provide building blocks for rapidly evolving genes with new recognition capabilities or act as reservoirs of unique specificities. These alternatives are not mutually exclusive. A resistance locus containing a large family of related genes could allow both independent evolution as well as periodic shuffling of domains through recombination or gene conversion. Not all genes may be active; function may be restored by infrequent recombination between active and inactive copies. In the human major histocompatability complex (MHC), the class I genes are present in ~20 copies, only three of which appear to be functional (Trowsdale, 1993). The additional copies appear to be pseudogenes; the sequence variation encoded in these pseudogenes may be recycled through recombination or gene conversion with functional genes. Many *RGC2* copies are expressed, although some are pseudogenes (Meyers et al., 1998). Therefore, it is possible that some *RGC2* members serve as a reservoir of diversity.

The Large, Duplicated Nature of Resistance Gene Clusters Has Several Evolutionary Consequences

The position and rate of recombination events influence the structure and function of the *RGC2* genes. Reciprocal recombination between alleles (or orthologs, in the case of genes introgressed from other species) alters sequences within individual family members, potentially changing the specificity of the gene. This type of recombination is probably important for generating variation in allelic series such as the *L* locus in flax. The repeated nature of the LRR region in resistance genes provides the possibility of unequal crossing over within alleles. This has been observed in mutants at the *M* locus in flax (Anderson et al., 1997) and *RPP5* in Arabidopsis (Parker et al., 1997); in addition, the size of the LRR region varies between alleles of *L* (Ellis et al., 1997). Moreover, *RGC2* genes also vary in the number of LRR repeats (Meyers et al., 1998). Unequal crossing over within the coding regions of paralogs generates chimeric genes; evidence of such events has been found at the *Xa21* locus of rice (Song et al., 1997). Furthermore, meiotic instability, alterations in resistance specificity, and the appearance of disease lesion mimics are associated with a high frequency of unequal crossovers at the *Rp1* locus in maize (Sudupak et al., 1993; Richter et al., 1995; Hu et al., 1996). Unequal crossover events in regions flanking genes alter the number of genes but do not generate new specificities. Recombination events in the flanking regions have been detected at the *Cf-4*/*9* cluster in tomato (Parniske et al., 1997). Hybridization of $AC15_{800}$ indicated that the copy number of *RGC2* genes varies greatly between genotypes (Anderson et al., 1996); however, this does not reveal whether unequal crossing over has occurred within or outside of *RGC2* sequences. Phylogenetic analysis of *RGC2* genes did indicate that unequal crossing over has infrequently generated chimeric genes during the evolution of this cluster.

A variety of genetic events may be important in the evolution of new resistance specificities in plants. The *RGC2* family provides evidence for both recombination and diversifying selection (this study; Meyers et al., 1998). These evolutionary mechanisms are critical components for producing and maintaining sequence diversity in other multigene families involved in nonhost recognition. Infrequent recombination within the mammalian *MHC* and *Ig* gene clusters has resulted in gene duplication and occasional loss of function or deletion of duplicated sequences (Hughes and Yeager, 1997; Nei et al., 1997). The lack of congruency between physical position in the *RGC2* locus and sequence similarity indicates that there has been a complex series of recombination events resulting in duplications and deletions as well as chimeric genes. Frequent recombination would decrease variation through concerted evolution, resulting in the homogenization of sequences due to gene conversion within a multigene family (Smith, 1973; Dover, 1982). There is no evidence for such a homogenization of sequences in the *RGC2* cluster; sequence variability in this multigene family is high. Therefore, rates of unequal crossing over and gene conversion must be below levels required for concerted evolution. Diversifying selection is a major force increasing variation in *MHC* and *Ig* genes (Hughes and Nei, 1988; Ota and Nei, 1994; Nei et al., 1997). Diversifying selection also seems to be important in increasing the sequence diversity of the *RGC2* family members and other resistance genes (Parniske et al., 1997; Meyers et al., 1998; Wang et al., 1998). The

combined effects of infrequent recombination and diversifying selection on individual genes may permit plants to evolve resistance to rapidly changing pathogens.

METHODS

Identification of RAPD and AFLP Markers

Two sets of lettuce (*Lactuca sativa*) genomic DNAs were screened to identify markers linked to *Dm* genes. The initial set of DNAs was composed of resistant and susceptible bulked DNAs for the *Dm1*/*3*, *Dm4*/*7*, and *Dm5*/*8* clusters as well as six deletion mutants representing *Dm1*, *Dm3*, *Dm5*/*8*, and *Dm7* (mutants were dm1b, dm1d, dm3r1208/dm3r1885 [these two DNAs were pooled to create a minimum deletion around *Dm3*], dm3r1608, dm58a, and dm7a; mutants described in Okubara et al., 1994). These DNAs were screened with 336 randomly amplified polymorphic DNA (RAPD) primers (Operon Technologies Inc., Alameda, CA; Williams et al., 1991) and 80 amplified polymorphism (AFLP) primer pairs with EcoRI and MseI adapters and three selective bases per primer (Keygene, Wageningen, The Netherlands; Vos et al., 1995). To increase the speed of marker screening, we reduced the set of DNAs to a panel of four deletion mutants (dm1d, dm3r1608, dm58a, and dm7a). These four DNAs were screened with 500 RAPD primers and 648 combinations of AFLP primer pairs (EcoRI [+4 selective bases] and MseI [+3 selective bases]).

Isolation of BAC Clones

Genomic bacterial artificial chromosome (BAC) libraries of lettuce had previously been constructed from cultivar Diana by using partial HindIII and EcoRI digests (Frijters et al., 1997; Z. Zhang and R.W. Michelmore, unpublished data). These libraries represent two to three genome equivalents in 76,000 clones with an average insert size of 115 kb. Clones were gridded onto Hybond N+ membranes (Amersham Corp., Arlington Heights, IL) with 1536 BACs per filter. Duplicate filters of both libraries were screened with a variety of probes: the cloned RAPD band $OPAC15_{800}$ (Anderson et al., 1996), the nucleotide binding site (NBS) of *RGC2B* (Shen et al., 1998), and $IPCR_{800}$ (Okubara et al., 1997). DNA gel blot hybridizations were performed according to standard protocols (Sambrook et al., 1989). Clones that hybridized strongly on the gridded filters were miniprepped (Sambrook et al., 1989), digested with HindIII, and rechecked with DNA gel blot hybridization by using an ECL (Amersham Corp.) chemiluminescence kit according to the supplier's instructions. Sizes of BAC clones were determined by NotI digests analyzed on pulsed-field gels.

Analysis of BAC Clones

BACs were assayed for the presence of markers that were missing in the largest deletion mutant, dm3r1608. These include AFLP markers B13CG01, B13CG25, B13AG05, S05AG02, S05TG43, S05TG44, and S05GF06; microsatellite marker MSAT15-34 (Okubara et al., 1997); sequence characterized amplified regions (SCAR) marker SCE14 (derived from OPE14 in Anderson et al., 1996); and restriction fragment length polymorphism (RFLP) markers $AC15_{800}$ (Anderson et al., 1996), AM14 (Anderson et al., 1996), K13 (Anderson et al., 1996), RGC2 (Shen et al., 1998), and $IPCR_{800}$ (Okubara et al., 1997).

AFLP fingerprints (Vos et al., 1995) were obtained for individual BAC clones by use of fluorescently labeled primers with bands sized on an ABI 377 sequencer (Applied Biosystems Inc., Foster City, CA). The primers used were C35, a nonselective EcoRI primer, and H18, H19, H20, and H21, which are MseI primers each with a single selective nucleotide (Table 1). Fingerprint patterns were detected and analyzed using GeneScan 2.0 (Applied Biosystems Inc.). Microsatellite markers were amplified under reaction conditions identical to those used for MSAT15-34 (Okubara et al., 1997). Polymerase chain reactions (PCRs) on BACs were performed using 5 ng of miniprepped plasmid DNA.

HindIII fingerprints were obtained for individual BAC clones by digestion of BAC DNA and analysis of the resulting fragment sizes. Five micrograms of miniprepped plasmid DNA was digested with HindIII, separated on 1% agarose, and transferred to a Hybond N+ membrane (Amersham Corp.) for DNA gel blot analysis. Fragments observed on the agarose gel were compared between BACs to identify duplicated bands as potentially overlapping fragments. In several cases, overlap was checked by DNA gel blot hybridization using radioactively labeled BAC DNA digested with HindIII.

Low-Copy-Number Subclones

Six libraries were constructed by partially digesting BACs H15, H2, and E32 with either Tsp509I or Sau3AI (New England Biolabs, Beverly, MA). Enzyme digests were done according to standard techniques (Sambrook et al., 1989), with limiting enzyme concentrations used to obtain partial digests. Fragments >3 kb were gel purified and cloned into either EcoRI- or BamHI-digested pUC119 (for Tsp509I or Sau3AI libraries, respectively). The majority of resulting transformants contained inserts of 1.5 to 2 kb. Transformants representing approximately three BAC equivalents for each digest of each BAC were screened by reverse genomic DNA gel blot hybridizations. Lettuce genomic DNA was labeled by the random hexamer method (Amersham Corp.) and used to probe the clones arrayed on membranes. Hybridization intensity was determined using a STORM PhosphorImager (Molecular Dynamics, Sunnyvale, CA). Colonies with hybridization intensities comparable to control single-copy clones were selected. Clones containing BAC vector sequences were identified by hybridizing with empty BAC vector and removed. Approximately 15 low-copy clones from each library (30 from each BAC) with inserts >0.5 kb were sequenced and analyzed by BLAST searches (Altschul et al., 1997).

PCR-Based BAC End Clone Rescue

Inverse PCR primers were designed with pairs of primers adjacent to either side of the BAC cloning site (Table 1). For BACs with inserts at the HindIII site (designated BAC H), inverse PCR templates were constructed from miniprepped BAC DNA and separately digested in 20-μL reactions with each of the four-base cutters MseI, HaeIII, and Tsp509I. Digested DNA was then diluted in 100 μL of T4 ligase buffer with 40 units of T4 ligase (New England Biolabs) and ligated overnight. Two microliters of this template was amplified using either the primer pairs ENDCLN1/ENDCLN2 or ENDCLN3/ENDCLN4 (Table 1)

to amplify each end of the BAC. In a similar manner, BACs with lettuce genomic inserts at the EcoRI cloning site (designated BAC E) were analyzed by use of inverse PCR templates constructed as above, using MseI, HaeIII, and Sau3AI digests and primer pairs ENDCLN1/ENDCLN4 or T7 (Promega, Madison, WI)/ENDCLN5 (Table 1) to amplify the left or right ends, respectively. PCR products were sequenced and analyzed by BLAST searches (Altschul et al., 1997).

PCR-Based Sequencing and Analysis

All PCRs were performed in 30 cycles with a 58 to 60°C annealing temperature. Half of the PCR product was checked on an agarose gel, and the other half was treated with exonuclease I and shrimp alkaline phosphatase (Amersham Corp.), according to the supplier's instructions. When possible, PCR products were then sequenced directly (see below). PCRs resulting in multiple bands or high background were cloned using the pGEMT vector system (Promega), according to the manufacturer's instructions; multiple clones were sequenced to eliminate PCR artifacts.

Sequencing reactions were performed with a dye terminator cycle sequencing kit (Applied Biosystems Inc.) and either the original PCR primers or standard Sp6, T7, M13 (-21), or M13 reverse primers (Promega). Reactions were then resolved on an ABI 377 automated sequencer (Applied Biosystems Inc.). Sequence data were evaluated using Sequencher (Gene Codes, Ann Arbor, MI) for sequence editing and contig assembly. The DNAstar (Lasergene, Madison, WI) and Genetics Computer Group (Madison, WI) software packages were used for multiple sequence alignments and sequence comparisons. Phylogenetic studies were performed using PAUP*, version 4.0 (Sinauer Associates, Sunderland, MA). BLAST searches (Altschul et al., 1997) were performed by using the National Center for Biotechnology Information (Bethesda, MD) website (www.ncbi.nlm.nih. gov). GenBank accession numbers for sequences used in the phylogenetic analysis are given in Table 2.

ACKNOWLEDGMENTS

We gratefully acknowledge the technical assistance of Pejman Rohani and Michelle Lai. B.C.M. was supported in part by a National Science Foundation graduate research fellowship. This project was supported by the U.S. Department of Agriculture National Research Initiative Competitive Grant Program (Grant No. 95-37300-1571).

Received June 12, 1998; accepted September 14, 1998.

REFERENCES

Aarts, M.G.M., te Lintel Hekkert, B., Holub, E.B., Beynon, J.L., Stiekema, W.J., and Periera, A. (1998). Identification of *R*-gene homologous DNA fragments genetically linked to disease resistance loci in *Arabidopsis thaliana.* Mol. Plant-Microbe Interact. **11,** 251–258.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res. **25,** 3389–3402.

Anderson, P.A., Okubara, P.A., Arroyo-Garcia, R., Meyers, B.C., and Michelmore, R.W. (1996). Molecular analysis of irradiation-induced and spontaneous deletion mutants at a disease resistance locus in *Lactuca sativa.* Mol. Gen. Genet. **251,** 316–325.

Anderson, P.A., Lawrence, G.J., Morrish, B.C., Ayliffe, M.A., Finnegan, E.J., and Ellis, J.G. (1997). Inactivation of the flax rust resistance gene *M* associated with loss of a repeated unit within the leucine-rich repeat coding region. Plant Cell **9,** 641–651.

Arumuganathan, K., and Earle, E.D. (1991). Nuclear DNA content of some important plant species. Plant Mol. Biol. Rep. **9,** 208–218.

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

Bevan, M., et al. (1998). Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana.* Nature **391,** 485–488.

Bonnier, J.F.M., Reinink, K., and Groenwald, R. (1994). Genetic analysis of *Lactuca* accession with the major gene resistance to lettuce downy mildew. Phytopathology **78,** 462–468.

Botella, M.A., Coleman, M.J., Hughes, D.E., Nishimura, M.T., Jones, J.D.G., and Somerville, S.C. (1997). Map positions of 47 *Arabidopsis* sequences with sequence similarity to disease resistance genes. Plant J. **12,** 1197–1211.

Crute, I.R., and Pink, D.A.C. (1996). Genetics and utilization of pathogen resistance in plants. Plant Cell **8,** 1747–1755.

Dixon, M.S., Jones, D.A., Keddie, J.S., Thomas, C.M., Harrison, K., and Jones, J.D.G. (1996). The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. Cell **84,** 451–459.

Dover, G.A. (1982). Molecular drive, a cohesive mode of species evolution. Nature **299,** 111–117.

Ellis, J., Lawrence, G., Ayliffe, M., Anderson, P., Collins, N., Finnegan, J., Frost, D., Luck, J., and Pryor, T. (1997). Advances in the molecular genetic analysis of the flax–flax rust interaction. Annu. Rev. Phytopathol. **35,** 271–291.

Farrara, B., Ilott, T.W., and Michelmore, R.W. (1987). Genetic analysis of factors for resistance to downy mildew (*Bremia lactucae*) in lettuce (*Lactuca sativa*). Plant Pathol. **36,** 499–514.

Fedoroff, N.V. (1989). Maize transposable elements. In Mobile DNA, D.E. Berg and M.M. Howe, eds (Washington, DC: American Society for Microbiology), pp. 375–411.

Flor, H.H. (1971). Current status of the gene-for-gene concept. Annu. Rev. Phytopathol. **9,** 275–296.

Frijters, A.C.J., Zhang, Z., van Damme, M., Wang, G.-L., Ronald, P.C., and Michelmore, R.W. (1997). Construction of a bacterial artificial chromosome library containing large EcoRI and HindIII genomic fragments of lettuce. Theor. Appl. Genet. **94,** 390–399.

Frohman, M.A., Dush, M.K., and Martin, G.R. (1988). Rapid production of full-length cDNAs from rare transcripts by amplification using a single gene-specific primer. Proc. Natl. Acad. Sci. USA **85,** 8998–9002.

Hammond-Kosack, K.E., and Jones, J.D.G. (1997). Plant disease resistance genes. Annu. Rev. Plant Physiol. Plant Mol. Biol. **48,** 575–607.

Holub, E.B. (1997). Organization of resistance genes in *Arabidopsis.* In The Gene-for-Gene Relationship in Host–Parasite Interactions, I.R. Crute, E.B. Holub, and J.J. Burdon, eds (New York: CAB International), pp. 5–26.

Hu, G., Richter, T.E., Hulbert, S.H., and Pryor, T. (1996). Disease lesion mimicry caused by mutations in the rust resistance gene *rp1.* Plant Cell **8,** 1367–1376.

Hughes, A.L., and Nei, M. (1988). Pattern of nucleotide substitution at major histocompatability complex class I loci reveals overdominant selection. Nature **335,** 167–170.

Hughes, A.L., and Yeager, M. (1997). Molecular evolution of the vertebrate immune system. Bioessays **19,** 777–786.

Hulbert, S.H. (1997). Structure and evolution of the *Rp1* complex conferring rust resistance in maize. Annu. Rev. Phytopathol. **35,** 293–310.

Ilott, T.W., Hulbert, S.H., and Michelmore, R.W. (1989). Genetic analysis of the gene-for-gene interaction between lettuce (*Lactuca sativa*) and *Bremia lactucae.* Phytopathology **79,** 888–897.

Islam, M.R., and Shepherd, K.W. (1991). Present status of genetics of rust resistance in flax. Euphytica **55,** 255–267.

Jones, D.A., Dickinson, M.J., Balint-Kurti, P.J., Dixon, M.S., and Jones, J.D.G. (1993). Two complex resistance loci revealed in tomato by classical and RFLP mapping of the *Cf-2*, *Cf-4*, *Cf-5*, and *Cf-9* genes for resistance to *Cladosporium fulvum.* Mol. Plant-Microbe Interact. **6,** 348–357.

Jorgensen, J.H. (1994). Genetics of powdery mildew resistance in barley. Crit. Rev. Plant Sci. **13,** 97–119.

Kanazin, V., Marek, L.F., and Shoemaker, R.C. (1996). Families of resistance gene analogs are conserved and clustered. Proc. Natl. Acad. Sci. USA **93,** 11746–11750.

Kesseli, R.V., Paran, I., and Michelmore, R.W. (1994). Analysis of a detailed genetic linkage map of *Lactuca sativa* (lettuce) constructed from RFLP and RAPD markers. Genetics **136,** 1435–1446.

Kunkel, B.N. (1996). A useful weed put to work—Genetic analysis of disease resistance in *Arabidopsis thaliana.* Trends Genet. **12,** 63–69.

Lagudah, E.S., Moullet, O., and Appels, R. (1997). Map-based cloning of a gene sequence encoding a nucleotide-binding domain and a leucine-rich region at the *Cre3* nematode resistance locus of wheat. Genome **40,** 659–665.

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

Maissonneuve, B., Bellec, Y., Anderson, P., and Michelmore, R.W. (1994). Rapid mapping of two genes for resistance to downy mildew from *Lactuca serriola* to existing clusters of resistance genes. Theor. Appl. Genet. **89,** 96–104.

Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganal, M.W., Spivey, R., Wu, T., Earle, E.D., and Tanksley, S.D. (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. Science **262,** 1432–1436.

Meyers, B.C. (1998). A Molecular Analysis of Clusters of Resistance Genes in Lettuce. PhD Dissertation (Davis, CA: University of California).

Meyers, B.C., Shen, K.A., Rohani, P., Gaut, B.S., and Michelmore, R.W. (1998). Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. Plant Cell **10,** 1833–1846.

Michealson, M.J., Price, H.J., Ellison, J.R., and Johnston, J.S. (1991). Comparison of plant DNA contents determined by fleugen microspectrophotometry and laser flow cytometry. Am. J. Bot. **78,** 183–188.

Michelmore, R.W., and Meyers, B.C. (1998). Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. Genome Res., in press.

Michelmore, R.W., Paran, I., and Kesseli, R.V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. Proc. Natl. Acad. Sci. USA **88,** 9828–9832.

Nei, M., Gu, X., and Sitnikova, T. (1997). Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proc. Natl. Acad. Sci. USA **94,** 7799–7806.

Okubara, P.A., Anderson, P.A., Ochoa, O.E., and Michelmore, R.W. (1994). Mutants of downy mildew resistance in *Lactuca sativa* (lettuce). Genetics **137,** 867–874.

Okubara, P.A., Arroyo-Garcia, R., Shen, K.A., Mazier, M., Kim, S.J., Yang, C.-H., and Michelmore, R.W. (1997). A transgenic mutant of *Lactuca sativa* (lettuce) with a T-DNA tightly linked to loss of downy mildew resistance. Mol. Plant-Microbe Interact. **10,** 970–977.

Ori, N., Eshed, Y., Paran, I., Presting, G., Aviv, D., Tanksley, S., Zamir, D., and Fluhr, R. (1997). 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 **9,** 521–532.

Ota, T., and Nei, M. (1994). Divergent evolution and evolution by the birth-and-death process in the immunoglobulin V_H gene family. Mol. Biol. Evol. **11,** 469–482.

Paran, I., and Michelmore, R.W. (1993). Development of reliable PCR-based markers linked to downy mildew resistance genes in lettuce. Theor. Appl. Genet. **85,** 985–993.

Parker, J.E., Coleman, M.J., Szabò, V., Frost, L.N., Schmidt, R., van der Biezen, E.A., Moores, T., Dean, C., Daniels, M.J., and Jones, J.D.G. (1997). The Arabidopsis downy mildew resistance gene *RPP5* shares similarity to the Toll and interleukin-1 receptors with *N* and *L6.* Plant Cell **9,** 879–894.

Parniske, M., Hammond-Kosack, K.E., Golstein, C., Thomas, C.M., Jones, D.A., Harrison, K., Wulff, B.B.H., and Jones, J.D.G. (1997). Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4*/*9* locus of tomato. Cell **91,** 821–832.

Pryor, T. (1987). The origin and structure of fungal disease resistance in plants. Trends Genet. **3,** 157–161.

Richter, T.E., Pryor, T.J., Bennetzen, J.L., and Hulbert, S.H. (1995). New rust resistance specificities associated with recombination in the *Rp1* complex in maize. Genetics **141,** 373–381.

Ronald, P.C., Albano, B., Tabien, R., Abenes, L., Wu, K.S., McCouch, S., and Tanksley, S.D. (1992). Genetic and physical analysis of the rice bacterial blight disease resistance locus, *Xa21.* Mol. Gen. Genet. **236,** 113–120.

Salmeron, J.M., Oldroyd, G.E.D., Rommens, C.M.T., Scofield, S.R., Kim, H.-S., Lavelle, D.T., Dahlbeck, D., and Staskawicz, B.J.

(1996). Tomato *Prf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. Cell **86,** 123–133.

Sambrook, J., Fritisch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

SanMiguel, P., Tikhonov, A., Jin, Y.K., Motchoulskaia, N., Zakharov, D., Melake-Berhan, A., Springer, P.S., Edwards, K.J., Lee, M., Avramova, Z., and Bennetzen, J.L. (1996). Nested retrotransposons in the intergenic regions of the maize genome. Science **274,** 765–768.

Saxena, K.M.S., and Hooker, A.L. (1968). On the structure of a gene for disease resistance in maize. Proc. Natl. Acad. Sci. USA **61,** 1300–1305.

Shen, K.A., Meyers, B.C., Islam-Faridi, N., Stelly, D.M., and Michelmore, R.W. (1998). Resistance gene candidates identified using PCR with degenerate oligonucleotide primers map to resistance gene clusters in lettuce. Mol. Plant-Microbe Interact. **11,** 815–823.

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

Smith, G.P. (1973). Unequal crossover and the evolution of multigene families. Cold Spring Harbor Symp. Quant. Biol. **38,** 507–513.

Song, W.-Y., Wang, G.-L., Chen, L., Kim, H.S., Pi, L.-Y., Holsten, T., Gardner, J., Wang, B., Zhai, W.X., Zhu, L.H., Fauquet, C., and Ronald, P.C. (1995). A receptor kinase–like protein encoded by the rice disease resistance gene, *Xa21.* Science **270,** 1804–1806.

Song, W.-Y., Pi, L.-Y., Wang, G.-L., Gardner, J., Holsten, T., and Ronald, P.C. (1997). Evolution of the rice *Xa21* disease resistance gene family. Plant Cell **9,** 1279–1287.

Sudupak, M.A., Bennetzen, J.L., and Hulbert, S.H. (1993). Unequal exchange and meiotic instability of disease-resistance genes in the *Rp1* region of maize. Genetics **133,** 119–125.

Trowsdale, J. (1993). Genomic structure and function in the MHC. Trends Genet. **9,** 117–122.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van der Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995). AFLP: A new technique for DNA fingerprinting. Nucleic Acids Res. **23,** 4407–4414.

Walker, E.L., Tobbins, T.P., Bureau, T.E., Kermicle, J., and Dellaporta, S.L. (1995). Transposon-mediated chromosomal rearrangements and gene duplications in the formation of the maize *R-r* complex. EMBO J. **14,** 2350–2363.

Wang, G.-L., Ruan, D.L., Song, W.-Y., Sideris, S., Chen, L., Pi, L.-Y., Zhang, S., Zhang, Z., Fauquet, C., Gaut, B.S., Whalen, M.C., and Ronald, P.C. (1998). *Xa21D* encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. Plant Cell **10,** 765–780.

Wessler, S.R., Bureau, T.E., and White, S.E. (1995). LTR-retrotransposons and MITEs: Important players in the evolution of plant genomes. Curr. Opin. Genet. Dev. **5,** 814–821.

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

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V. (1991). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. **18,** 6531–6535.

Witsenboer, H., Kesseli, R.V., Fortin, M.G., Stanghellini, M., and Michelmore, R.W. (1995). Sources and genetic structure of a cluster of genes for resistance to three pathogens in lettuce. Theor. Appl. Genet. **91,** 178–188.

Yu, Y.G., Buss, G.R., and Maroof, M.A.S. (1996). Isolation of a superfamily of candidate resistance genes in soybean based on a conserved nucleotide-binding site. Proc. Natl. Acad. Sci. USA **93,** 11751–11756.