

Inactivation of the Flax Rust Resistance Gene *M* Associated with Loss of a Repeated Unit within the Leucine-Rich Repeat Coding Region

Peter A. Anderson,^{a,1} Gregory J. Lawrence,^{a,b} Bronwyn C. Morrish,^a Michael A. Ayliffe,^b E. Jean Finnegan,^{a,b} and Jeffrey G. Ellis^{a,b,2}

^aCommonwealth Scientific and Industrial Research Organization, Division of Plant Industry, GPO Box 1600, Canberra, A.C.T., 2601, Australia

^bCooperative Research Centre for Plant Science, GPO Box 475, Canberra, A.C.T., 2601, Australia

The *M* rust resistance gene from flax was cloned after two separate approaches, an analysis of spontaneous *M* mutants with an *L6* gene-derived DNA probe and tagging with the maize transposon *Activator*, independently identified the same gene. The gene encodes a protein of the nucleotide binding site leucine-rich repeat class and is related (86% nucleotide identity) to the unlinked *L6* rust resistance gene. In contrast to the *L* locus, which contains a single gene with multiple alleles, ~15 related genes occur at the complex *M* locus, with only one encoding the *M* resistance specificity. The *M* protein contains two direct repeats of 147 and 149 amino acids in the C-terminal part of the leucine-rich region. Three mutant alleles of *M* encoding a product containing a single repeat unit of 154 amino acids were isolated. The mutant DNA sequences probably occurred by unequal intragenic exchange in the coding region of the repeats. The recombinant alleles lost *M* resistance and gained no detectable new resistance specificity.

INTRODUCTION

Genetic studies have shown that many different plant disease resistance genes occur in clusters and that two distinct arrangements may exist: either single genes with multiple alleles encoding different resistance specificities or a series of tightly linked genes forming complex loci (see Pryor and Ellis, 1993). Now that several resistance genes have been cloned (see Staskawicz et al., 1995; Boyes et al., 1996), analysis of the structural organization of these loci is being performed. For example, in tomato, the leaf mold *Cladosporium fulvum* resistance loci *Cf-9* (Jones et al., 1994) and *Cf-2* (Dixon et al., 1996) contain arrays of five or more related genes. In the case of *Cf-2*, two nearly identical genes in the cluster both encode *Cf-2* resistance specificity. Although only a single resistance specificity to tobacco mosaic virus (TMV) has been identified at the *N* locus in tobacco, molecular analysis has demonstrated that this locus is also an array of related genes, with only one controlling TMV resistance (Whitham et al., 1994). Another example of this arrangement is the *Pto* locus of tomato, which contains approximately five genes encoding serine/threonine kinase-like genes (Martin et al.,

1993, 1994). Only one of these genes encodes resistance to strains of *Pseudomonas syringae* pv *tomato* carrying the *Avr-Pto* avirulence gene (Martin et al., 1994). Interestingly, a second gene, *Prf*, necessary for *Pto*-mediated resistance but unrelated to *Pto*, occurs within this complex (Salmeron et al., 1996).

In contrast to these complex loci are the *RPS2* (Bent et al., 1994; Mindrinos et al., 1994) and *RPM1* (Grant et al., 1995) bacterial resistance genes from Arabidopsis. Both occur as apparently single-copy genes at their respective loci. Two naturally occurring alleles of *RPS2* have been identified: one is active in resistance, and the other confers no detectable resistance specificity. The *RPM1* gene consists of a single allele, and lines that do not express *RPM1* resistance lack the gene.

In flax (*Linum usitatissimum*), 31 genes that confer resistance to the fungal pathogen responsible for rust disease of flax (*Melampsora lini*) occur at five loci designated *K*, *L*, *M*, *N*, and *P*. These include both simple and complex loci (see Islam and Shepherd, 1991). The *L* locus with 13 resistance specificities and the *M* locus with seven are the most extensively analyzed. Genetic analyses indicate that genes in the *L* group are most likely allelic to each other (alternative forms of a single gene); no recombinants expressing both parental specificities have been detected between many pairs of *L* specificities in large numbers of testcross progeny

¹Current address: Flinders University of South Australia, School of Biological Sciences, GPO Box 2100, Adelaide, S.A., 5001, Australia.

²To whom correspondence should be addressed. E-mail ellisj@pican.pi.csiro.au; fax 61-6-246-5000.

(Flor, 1965; Shepherd and Mayo, 1972; Islam et al., 1989). In contrast, similar studies involving pairs of *M* specificities have recovered recombinants expressing both parental specificities, which indicates that at least some members of the *M* group are separate but closely linked genes (Flor, 1965; Mayo and Shepherd, 1980).

After the recent cloning of the *L6* allele (Lawrence et al., 1995), Ellis et al. (1995) found that *L6*-derived DNA probes hybridized with a family of related restriction fragments in flax DNA. Depending on the restriction enzyme used, up to 15 such fragments could be identified by gel blot hybridization. With certain restriction enzymes, some fragments from the line Forge (which contains four resistance genes—*L6*, *M*, *N*, and *P2*) were polymorphic with respect to fragments from the line Hoshangabad (which contains no known functional rust resistance genes). In a testcross family, (Forge × Hoshangabad) × Hoshangabad, these polymorphic fragments were found to cosegregate exactly with either the *L6* or the unlinked *M* gene, indicating that genes at the *M* locus have similarities to those at the *L* locus (Ellis et al., 1995). Furthermore, whereas several fragments cosegregated with the *M* gene, only a single fragment was ever identified that cosegregated with the *L6* gene. These observations pointed to there being a cluster of related genes at the *M* locus but most likely only a single gene at the *L* locus. This finding is fully consistent with the earlier genetic fine-structure studies. This work could not map the multiple DNA fragments common (monomorphic) to both Forge and Hoshangabad that did not segregate in the testcross. In this study, we describe the use of pulsed-field gel electrophoresis of high molecular weight DNA to demonstrate that the family of *L6*-related fragments occurs at only two locations, the *L* and *M* loci. Consequently, if the *L* locus contains only a single gene, it can be estimated that the array of related genes at the *M* locus contains ~15 members.

The above-mentioned observations indicated that *L6*-derived probes could be used to access genes at the *M* locus, but because this locus contains an array of related genes, the observations also raised the problem of how to identify which particular array member in the line Forge was responsible for the *M* rust resistance specificity. In this study, we report the use of two separate approaches, analysis of spontaneous *M* mutants with an *L6*-derived DNA probe and *Ac* (for *Activator*) tagging, to independently identify the *M* gene.

The *M* gene, like *L6*, is predicted to encode a product of the nucleotide binding site, leucine-rich repeat (NBS-LRR) class of disease resistance proteins (see Staskawicz et al., 1995; Boyes et al., 1996). Thus, this class of resistance genes occurs at two loci with contrasting physical structures. The LRR motif, frequently found in proteins involved in protein-protein interactions (Kobe and Deisenhoffer, 1994), occurs in a wide range of disease resistance genes (Jones and Jones, 1997). It has been proposed that variation in the LRR sequences may provide the discriminatory capacity for the specific recognition functions of gene-for-gene disease resistance (Staskawicz et al., 1995). Our experiments involving

domain swaps between alleles of the *L* locus are providing support for this view (J.G. Ellis, G.J. Lawrence, and K.W. Shepherd, unpublished results). Hence, mechanisms capable of generating variation in the LRR region are of particular interest. In this study, we describe three independent but identical LRR region mutants of the *M* gene: whereas the wild-type gene contains two direct DNA repeats encoding 147 and 149 amino acids, the mutants contain a single repeat. The three mutant alleles could be a consequence of unequal exchange between the first repeat in one *M* gene and the second repeat in its homolog. Thus, the DNA repeats can participate in events leading to structural alteration of resistance genes and may play a significant role in the evolution of rust resistance genes. These mutant alleles of the *M* gene lost *M* resistance specificity and gained no detectable new resistance specificity.

RESULTS

Strategy for Isolating the *M* Rust Resistance Gene

Two approaches were used to identify the *M* resistance gene. The first involved an examination of 41 independent *M* gene mutants by gel blot hybridization using an *L6*-derived probe. In the second, the *M* gene was tagged with the maize transposable element *Ac*. Both approaches pinpointed the same gene.

The 41 *M* gene mutants were identified as rust-susceptible individuals among 276,000 F_1 progeny derived from Forge × Hoshangabad crosses that were produced in a program designed to tag the four resistance genes in the Forge parent (*L6*, *M*, *N*, and *P2*) with either the *Ac* or *Tam3* transposons. Three broad categories of Forge parents were used to make the F_1 progeny. The first contained either no introduced transposons or one or a few copies of *Ac* or *Tam3* in which no transposon activity was detected. The second contained eight to 15 active copies of *Ac*. None was linked to *M*. The third category contained the same multiple copies of *Ac* and in addition an extra *Ac* copy on a T-DNA inserted 1 map unit from *M*. The frequency of *M* mutants was 20 from 180,000 progeny (one in 9000) for the first category, four from 46,000 progeny (one in 11,000) for the second, and 17 from 50,000 (one in 3000) from the third category, which contained the *M*-linked *Ac*. One of these 17 mutants contained an *Ac*-tagged *M* gene (see below). This effect of the linked *Ac* on the mutation frequency of *M* is similar to our previous observation in which *Ac* linked to *L6* increased this gene's mutation frequency (Lawrence et al., 1993), probably due to a high frequency of deletion events resulting from *Ac* excision-induced chromosome breakage.

To eliminate large deletions from further analysis, all 41 *M* gene mutants were screened with two DNA probes that identify two restriction fragment length polymorphism markers (LU-2-2A and X22A-1A), closely linked to and flanking

the *M* locus (Ellis et al., 1995). Twenty mutants, including 12 from parents containing the *M*-linked *Ac*, had lost one or both flanking markers: these were not analyzed further because it was likely that they contained large deletions at the *M* locus. The remaining mutants were analyzed for the presence of newly transposed *Ac* elements linked to *M*, as previously described (Lawrence et al., 1993, 1995). In addition, the gel blot hybridization patterns of these mutants were examined using probe Lu-3 derived from the LRR-encoding region of the *L6* gene (Ellis et al., 1995) to determine whether any of them might reveal evidence of a localized mutation event at the *M* locus, such as loss or alteration of a single Forge fragment.

Mutants X32, X39, and X123 Contain Identical DNA Alterations in an *L6*-like Gene at the *M* Locus

Gel blot hybridization of DNA from the mutants, using probe Lu-3, identified three mutants (X32, X39, and X123) that contained a novel *Xba*I restriction fragment not present in the parents. This novel fragment was identical in size in all three mutants (Figure 1A). This difference was more clearly evident when DNA was digested with *Sal*I, which produced a less complex fragment pattern (Figure 1B). This *Sal*I pattern was of particular interest because not only did the three mutants each contain a novel 3.0-kb fragment but they all lacked a 3.5-kb fragment present in the Forge parent (Figure

1B), suggesting that the novel 3.0-kb fragment could be an altered form of the parental 3.5-kb fragment.

Evidence for an *M* locus origin of this fragment came from the observation that a near-isogenic line of Bison, developed by H.H. Flor (Flor, 1954), containing the *M* gene from Dakota introgressed into Bison by 12 backcrosses (see Islam and Mayo, 1990), possessed the 3.5-kb fragment, whereas Bison lacked this fragment (Figure 1C). The *M* gene in Forge originated from Dakota.

Thus, in three independent mutants, loss of the *M* specificity was associated with an identical reduction in size of the 3.5-kb *M*-linked DNA fragment, suggesting that this DNA fragment was derived from the *M* gene. The 3.5-kb *Sal*I fragment was cloned from the *M* parent, and partial sequence analysis confirmed that one end of the fragment was similar to the 3' end of the *L6* gene, starting at base pair 3389 (Lawrence et al., 1995) and continuing beyond the 3' end of *L6*. A 700-bp *Xba*I subfragment (probe Mxb-1) from this downstream region hybridized only with the 3.5-kb *Sal*I fragment in genomic DNA gel blots and provided an *M* locus-specific probe.

Probe Mxb-1 was used to screen a genomic library of DNA from Forge. One clone that hybridized with Mxb-1 and to probes from the 5' and 3' coding regions of *L6* was selected. This clone contained the 3.5-kb *Sal*I fragment originally identified in the *M* gene region (Figure 2A, map i). From this clone, an 11-kb *Eco*RI fragment containing the candidate *M* gene was subcloned for further analysis.

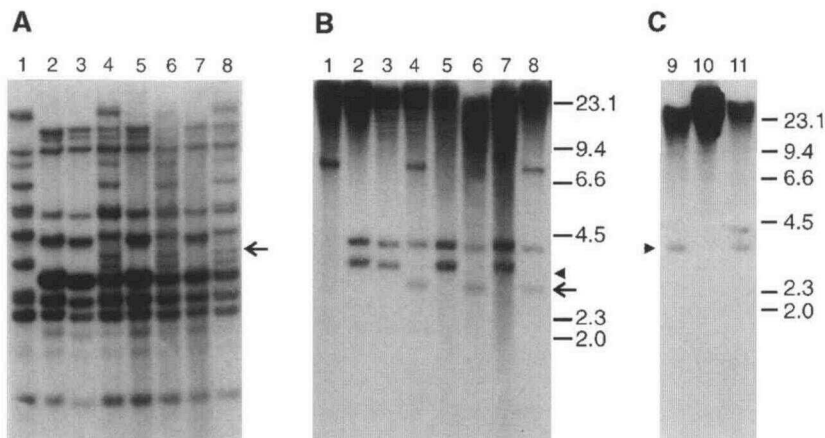


Figure 1. DNA Gel Blot Analysis of Parental Flax Lines and *M* Gene Mutants.

DNA was digested with *Xba*I and *Sal*I and hybridized with the Lu-3 probe from the LRR region of the *L6* gene. The mutants from which DNA was isolated are F_1 hybrids between Forge and Hoshangabad; thus, their expected restriction pattern is the sum of the fragments observed in the parents. (A) and (B) Lanes 1 contain Hoshangabad; lanes 2, Forge; lanes 3, Forge parent of X32 mutant; lanes 4, X32 mutant; lanes 5, Forge parent of X39 mutant; lanes 6, X39 mutant; lanes 7, Forge parent of X123 mutant; and lanes 8, X123 mutant. DNA was digested with *Xba*I in (A) and *Sal*I in (B). (C) Lane 9 contains the near-isogenic line of Bison containing the *M* gene from Dakota introduced by 12 backcrosses; lane 10, Bison; and lane 11, Dakota. DNA was digested with *Sal*I.

Arrowheads indicate the 3.5-kb *Sal*I fragment in the wild type, and arrows indicate altered fragments present in the mutant and absent in the Forge parents. Length markers in kilobases are indicated at right in (B) and (C).

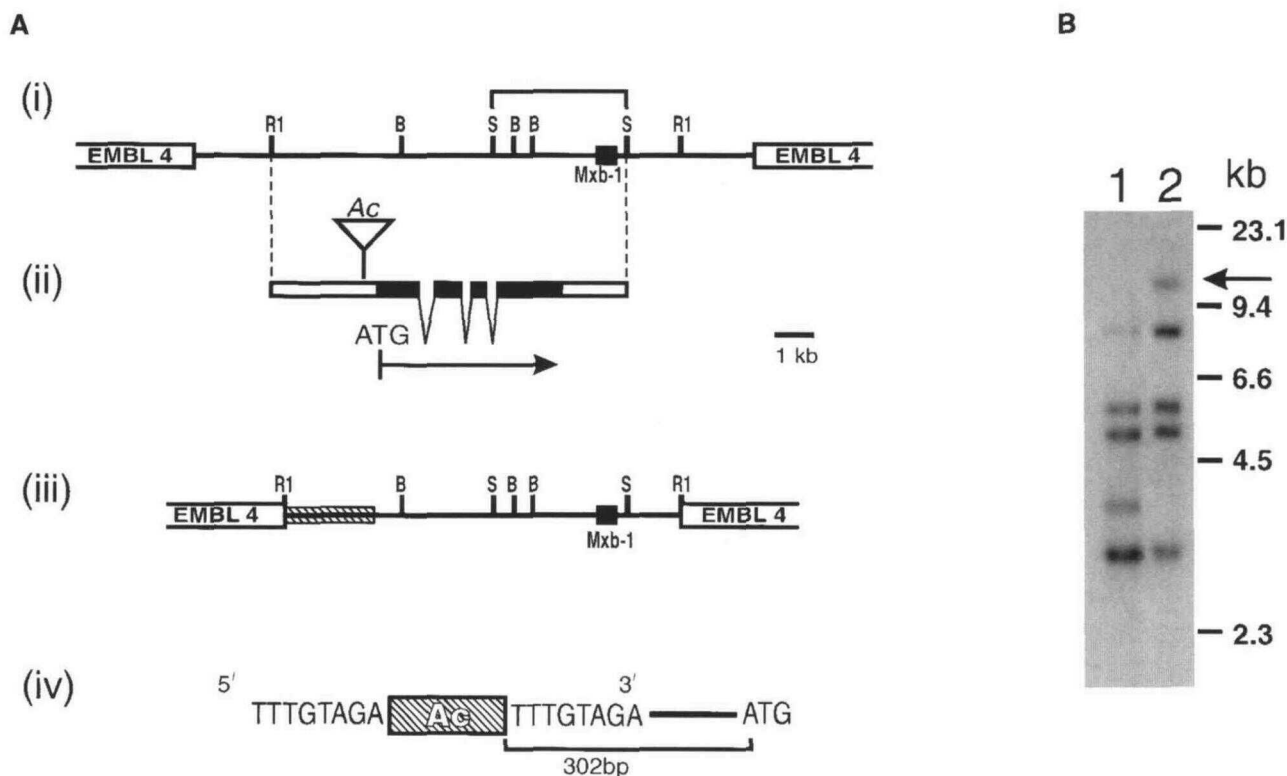


Figure 2. Maps of *M* Gene Clones, Location of Transposed *Ac*, and DNA Blot Showing Newly Transposed *Ac* in Mutant X142.

(A) A restriction map of the EMBL4 clone obtained from a Forge genomic library by hybridization with the *M*-specific Mxb-1 probe (indicated by the filled box) is shown in map i. The clone contains the 3.5-kb Sall fragment (see Figure 1B), indicated by the square bracket, that was altered in the three *M* gene mutants X32, X39, and X123. B, BamHI; R1, EcoRI; S, Sall. A 9.3-kb region of the EMBL4 clone was sequenced and is shown in map ii. This region contains the *M* gene. The position of the *Ac* element in the mutant X142 is indicated by the triangle. The positions of the three introns, which are represented by the saw-toothed line, were determined from the sequence of a partial cDNA (cDNA23.1) and by reverse transcriptase-polymerase chain reaction. The coding region beginning with ATG is indicated by the arrow. The EMBL4 clone of the EcoRI fragment containing the 3' *Ac*-plant DNA junction from the *M* gene mutant X142 is shown in map iii, with the *Ac* DNA indicated as a cross-hatched box from the internal EcoRI site to the 3' end of *Ac* and the position of Mxb-1 shown by the filled box. The 8-bp sequence duplication at the site of insertion of *Ac* in the mutant X142, which is 302 bp upstream of the ATG codon of the *M* gene, is shown in map iv.

(B) Shown is a DNA gel blot of genomic DNA from the *M* gene mutant X142 (lane 2) and its Forge parent (lane 1) digested with EcoRI and hybridized with the 3' end probe of the *Ac* element. The arrow indicates the new fragment in the X142 mutant that was cloned in EMBL4 (map iii in [A]) and that resulted from *Ac* transposition into the *M* gene. The length markers in kilobases are indicated at right.

Confirmation of the *M* Gene by *Ac* Tagging

One of the 17 *M* mutants recovered among Forge × Hoshangabad progeny in which the Forge parent possessed the *M*-linked *Ac* was unusual in that it had a partially susceptible phenotype (pinhead-sized pustules on the younger leaves) to rust strain CH5F2-78 that recognizes the *M* gene, in contrast to the remaining 40 mutants that were fully susceptible. The parent plants Forge and Hoshangabad were fully resistant (no spore production) and fully susceptible, respectively, to strain CH5F2-78. This unusual mutant (X142) contained a newly transposed *Ac* that was detected by gel blot analysis using a DNA probe derived from the 3' end of

Ac (Figure 2B). X142 was selfed, and 36 progeny were tested for their reaction to rust strain CH5F2-78 and assayed for the presence of the newly transposed *Ac* by gel blot analysis. The partially susceptible phenotype and the presence of the transposed *Ac* showed complete cosegregation; the transposed *Ac* was present in 29 partially susceptible progeny and absent in the seven fully susceptible progeny.

The transposed *Ac* in X142, together with flanking flax DNA, was cloned in two fragments from the internal EcoRI site of *Ac*. One clone contained an EcoRI fragment covering the 5' *Ac*-plant DNA junction, and the second contained the 3' *Ac*-plant DNA junction. The *M*-specific probe Mxb-1 hy-

bridized with the 3' *Ac*-plant junction clone illustrated in Figure 2A, map iii. Thus, the transposed *Ac* element in X142 was inserted into the same 11-kb *EcoRI* fragment that was identified above in the analysis of the spontaneous mutants X32, X39, and X123 and that was considered likely to contain the *M* gene.

Complementation Analysis Demonstrates That the Cloned Gene Controls *M* Resistance Specificity

Three different lines of flax were transformed with the 11-kb *EcoRI* fragment. One transformant of a line of flax homozygous for the mutant *M* allele *m-x32* was obtained and was resistant to rust strain CH5F2-78 that recognizes the *M* resistance gene but not the other resistance genes present in the Forge parent of X32 (*L6*, *N*, and *P2*). Two independent transformants of the flax line Cass also expressed the *M* resistance specificity. Hoshangabad, which contains no known rust resistance genes, was also transformed with the 11-kb *EcoRI* fragment. The single transformant from this experiment was resistant to CH5F2-78. Twenty-one progeny derived by selfing this plant were analyzed for rust resistance and for the presence of the transgene by DNA gel blot analysis. The transgene cosegregated with rust resistance; it was present in the 15 rust-resistant progeny and absent in the six rust-susceptible progeny. To confirm that the transgene expresses the *M* resistance specificity, one of the resistant progeny plants was inoculated with a second rust strain that does not recognize *M* but that does recognize *L6*, *N*, and *P2*. The plant was fully susceptible to this strain.

Structure of the *M* Gene

A 9.3-kb region derived from the genomic clone containing the *M* gene (Figure 2A, map i) was sequenced (GenBank accession number U73916) and compared with the *L6* gene. The *M* gene was 86% identical to the *L6* coding and intron sequences. The location of intron 1 (275 bases) in the *M* gene was determined by sequencing an incomplete cDNA clone (FC23.1) that spanned the 5' region of *M*, including intron 1. The positions of introns 2 and 3 were predicted by sequence alignment with the *L6* coding region and confirmed by a polymerase chain reaction (PCR) approach. Primers were made to sequences just upstream of the predicted ATG and just downstream of the translation stop codon, and long-range PCR was used to amplify a nearly full-length copy of the *M* transcript from cDNA. The 4.2-kb product of amplification was digested with *Xba*I and *Sac*I, whose sites occur in the *M* gene upstream of intron 2 and downstream of intron 3, respectively, and the resulting 1.8-kb fragment was cloned into a plasmid vector. Several clones were sequenced and were identical to the predicted coding region of *M*, indicating that the PCR reaction was specific for cDNA generated from *M* gene transcripts. Among these

clones, three classes were identified. One had introns 2 (75 bases) and 3 (94 bases) spliced as predicted from previous knowledge of *L6*, and the other two represented alternative transcripts involving intron 2. These transcripts, which would encode truncated products, will be described in detail in a later paper. Removal of the three intron sequences from the genomic sequence gave rise to a single open reading frame of 3915 bases, starting from an ATG codon predicted by comparison to *L6*. Stop codons occur in all frames upstream of the open reading frame.

Two direct repeats of 441 and 447 bp occur in the 3' half of the *M* gene. These repeats are 90% identical in their DNA sequence. Similar repeats occur in the same region of *L6* (Lawrence et al., 1995).

The DNA surrounding the *Ac* insertion in the *m-x142* allele was also cloned and sequenced. *Ac* had inserted 302 bp upstream of the ATG codon of *M* (Figure 2A, map iv) and was flanked by an 8-bp duplication of flax DNA. This suggests that the partially susceptible phenotype of plants carrying this insertion allele is probably the result of a reduction in promoter activity and consequent reduced amount of the resistance gene product.

Comparison of the *L6* and *M* Protein Sequences

Translation of the open reading frame from the *M* gene predicts a polypeptide product of 1305 amino acids (Figure 3). This sequence is 78% identical and 85% similar to *L6* and contains all of the motifs identified in *L6* (Lawrence et al., 1995), including the hydrophobic N terminus, the nucleotide binding domain, and two direct repeats of 147 and 149 amino acid residues in the leucine-rich C-terminal region. In the *M* gene product, these repeats share 77% amino acid identity. Substantial differences between *L6* and *M* occur in the 37-amino acid region just downstream of the conserved hydrophobic N-terminal sequence. The predicted amino acid sequences of *M* and *L6* are aligned in Figure 3.

Loss of a Repeated Unit in Mutants X32, X39, and X123 Is Associated with Loss of *M* Specificity

The novel 3.0-kb *Sal*I fragments in the mutant alleles *m-x32*, *m-x39*, and *m-x123* (Figure 1B) were cloned, and the *m-x39* fragment was completely sequenced. Comparison of the sequences of the *m-x39* mutant 3.0-kb fragment with the 3.5-kb fragment from the wild-type gene showed that the two fragments were identical, except for a deletion of 426 bp in the shorter fragment. The deletion had occurred within the region of the gene containing the two tandem repeats located in the 3' end of *M*. The deletion spanned the end of repeat 1 and the beginning of repeat 2. Consequently, whereas the wild-type gene contained two tandem repeats of 441 and 447 bp, the mutant allele *m-x39* contained only one copy of a 462-bp repeat unit, as illustrated in Figure 4A

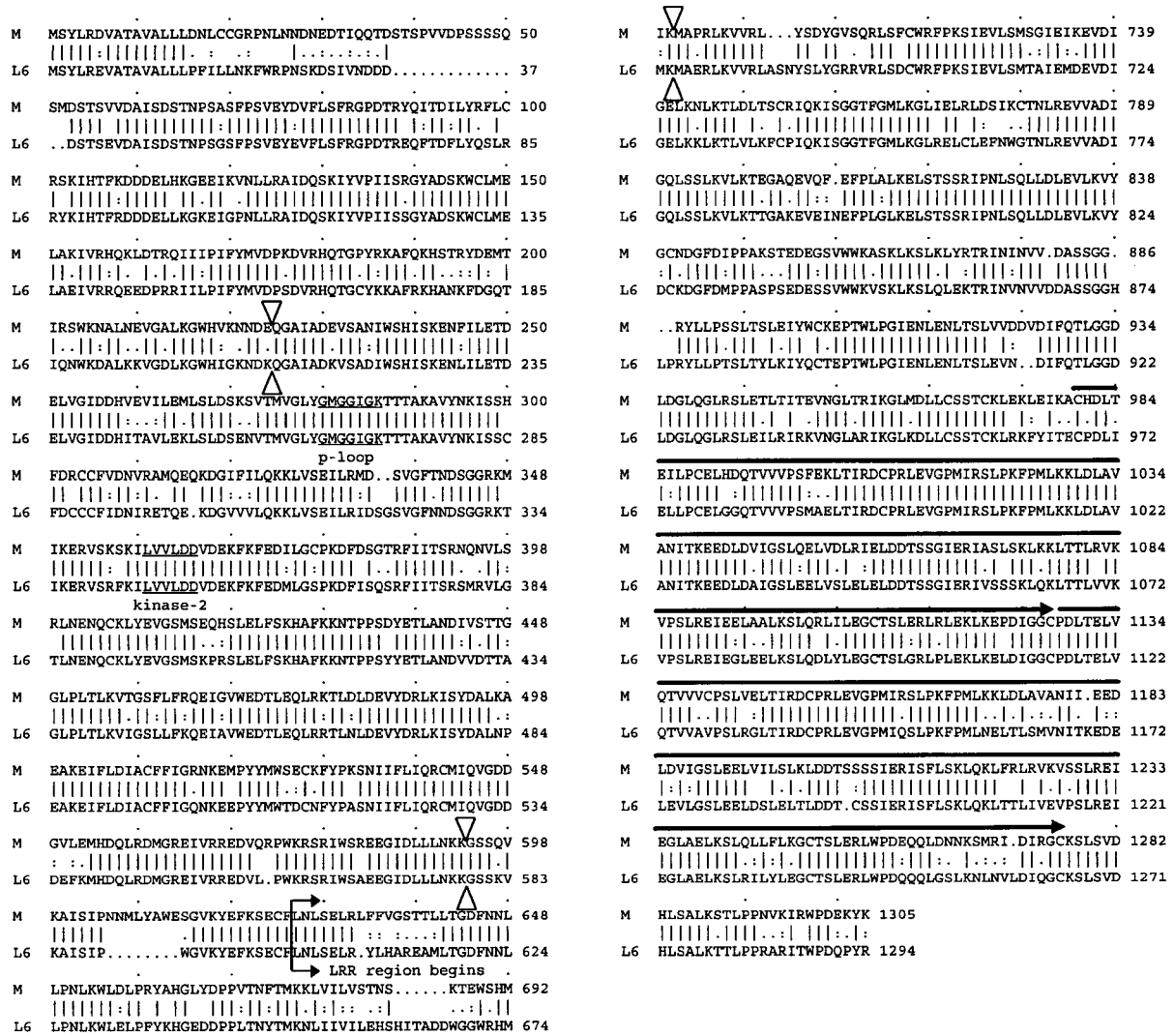


Figure 3. Comparison of the Predicted Amino Acid Sequence of L6 and M.

The L6 and M amino acid sequences were aligned using the sequence comparison program BESTFIT (Devereux et al., 1984). The P loop and kinase 2 domains are underlined, the beginning of the LRR region is indicated by the vertical line linking two short arrows, and the two large direct repeats that occur within the LRR region are indicated by long arrows in the C-terminal region of the protein. Open triangles indicate the position of the three introns in L6 and M. Gaps have been introduced to maximize alignment, vertical bars indicate identical amino acids, and the colons and dots represent similar amino acids.

(GenBank accession number U76370). Sequencing in the repeat region in *m-x32* and *m-x123* (data not shown) revealed that the three alleles contained identical changes.

Because the original reading frame is maintained in the mutant alleles, their predicted products are internally truncated M proteins containing a single repeat unit of 154 amino acids. The amino acid sequences of the two repeat regions of M and single repeat unit of *m-x39* are aligned in Figure 4B. The extra six amino acids at the beginning of repeat 1 and the extra eight amino acids at the end of repeat

2, indicated by boxes above the sequences in Figure 4B, provide distinguishing markers for the two repeats. Both the six- and eight-amino acid blocks are present in the single repeat unit in the mutants.

A related gene of unknown function, located in the M gene complex of Forge, also contains a single repeat unit. This gene, represented by the cDNA clone FC4 (Ellis et al., 1995; Lawrence et al., 1995), also encodes a single repeat unit containing both the six- and eight-amino acid blocks (Figure 4B).

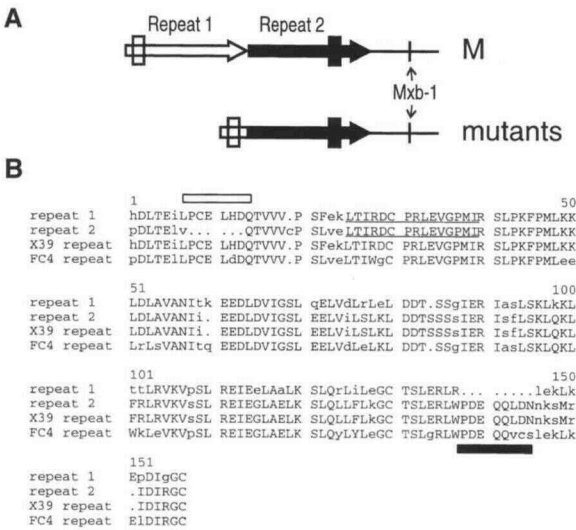


Figure 4. The Two Direct Repeats of the M Gene and the Single Repeat Units Found in *m-x39* and the Related cDNA FC4.

(A) Relationship of the two direct DNA repeats in M and the single repeat unit in mutant alleles *m-x32*, *m-x39*, and *m-x123*. The open and filled boxes show the location of extra bases at the beginning of repeat 1 and the end of repeat 2. The location of the M gene-specific sequence in probe Mxb-1 is indicated.

(B) Comparison of the predicted amino acid sequences of the repeats in the mutant allele *m-x39* (X39) and cDNA clone FC4. The open and filled boxes show the locations of extra amino acids at the beginning of repeat 1 and the end of repeat 2. The underlined sequences in repeats 1 and 2 are encoded by an interval of 45 invariant nucleotides in which intragenic exchange may have occurred in the formation of the three M gene mutants. Alignments were made using the PILEUP and PRETTY programs (Devereux et al., 1984). Gaps, indicated by dots, have been introduced to maximize the alignment. Dissimilar amino acids are indicated in lower case letters.

The mutant M allele *m-x32* was tested for new resistance specificities that may have arisen as a result of the intragenic exchange. A line homozygous for *m-x32* and not containing the L6, N, and P2 genes was isolated as described in Methods. This line was inoculated with a strain of rust isolated from the Australian native flax species *L. marginale*. This rust strain, which contains avirulence genes corresponding to most of the 31 resistance genes in cultivated flax, grew on the *m-x32* mutant. Thus, no new rust resistance specificity was detected.

The L6-Related Multigene Family Occurs at Only Two Loci in Flax

DNA probes from the coding region of the L6 gene detect a complex hybridization pattern on blots of genomic DNA di-

gested with six-base recognition site restriction enzymes (e.g., see Figure 1A). The question arises as to whether these fragments originate from only the L and M loci or whether additional loci are involved. To examine this question, high-molecular-weight BssHII restriction fragments from DNA derived from Forge and a line designated Trac1 were fractionated by pulsed-field gel electrophoresis and probed with the L6-derived probe Lu-3. As shown in Figure 5A, probe Lu-3 hybridized with only two fragments, of ~1000 and 200 kb, indicating that the L6-related fragments are derived from only two loci. To determine which of these fragments contained the L locus, we removed the Lu-3 probe, and the DNA blot was rehybridized with a probe from the 3' end of Ac. This probe hybridized only with the 200-kb fragment in the Trac1 line, which is homozygous for a single Ac inserted in the L6 gene (Figure 5B), thereby indicating that the 200-kb fragment contains the L locus. Therefore, the larger fragment should contain the M locus. This was confirmed (Figure 5C) with the M gene-specific probe Mxb-1. Similar results were observed in experiments in which the DNA was digested with NotI.

These experiments therefore demonstrate that genes closely related to the L6-derived probe, Lu-3, occur at only the L and M loci. Furthermore, the stronger intensity of hybridization of probe Lu-3 (Figure 5A) to the M locus fragment supports the earlier interpretation (Ellis et al., 1995) that the M locus is composed of a cluster of related genes. Based on the number of fragments detected in XbaI-digested DNA hybridized with probe Lu-3 from the L6 gene (Figure 1A), it is estimated that the M locus contains ~15 copies of L6-related genes.

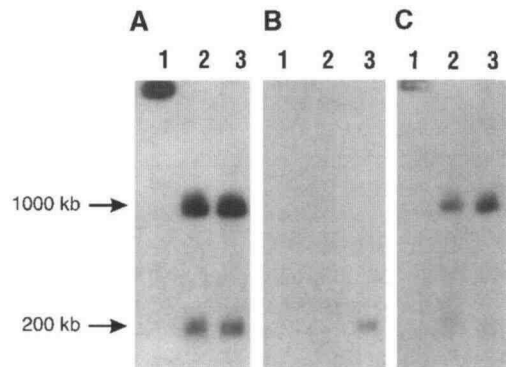


Figure 5. DNA Gel Blot of High Molecular Weight DNA from Flax Cut with BssHII and Separated by Pulsed-Field Gel Electrophoresis.

(A) Blot probed with Lu-3 from the LRR region of the L6 gene. **(B)** Blot probed with the 3' end of the Ac element. **(C)** Blot probed with the M gene-specific probe Mxb-1. Lanes 1 contain undigested DNA from Forge; lanes 2, digested DNA from Forge; and lanes 3, digested Trac1 DNA (homozygous for Ac in the L6 gene).

DISCUSSION

Structure of the *M* Locus

The *M* rust resistance locus in the flax line Forge contains an array of genes closely related to the unlinked *L6* gene. The *M* locus is <1000 kb and is estimated to contain ~15 related genes. Analysis of rust-susceptible mutants derived from Forge indicates that only one of the genes at the *M* locus, the designated *M* gene, is required for the *M* rust resistance specificity. The functions of the remaining, related genes at the locus, at least some of which are transcribed (Lawrence et al., 1995), are unknown. They may provide resistance to unidentified rust strains or to different types of pathogens. They may also act as a repository for sequence information for the generation of new resistance specificities by intergenic recombination.

Six other distinct rust resistance specificities, *M1* through *M6*, have been identified at the *M* locus in other lines of flax (see Islam and Shepherd, 1991). Gel blot analysis of these lines by using the Lu-3 probe shows that these specificities also occur in a complex locus of *L6*-related genes (P.A. Anderson and J.G. Ellis, unpublished results). Some of these specificities (*M1*, *M3*, and *M4*) have been recombined with the *M* resistance specificity and brought together on the same chromosome (Flor, 1965; Shepherd and Mayo, 1972; Mayo and Shepherd, 1980). In these cases, the recombinant *M* locus contains at least two genes that provide rust resistance of different specificity. It is possible that some of the other *L6*-related genes at the *M* locus of Forge are null alleles of these other active resistance genes.

The Cloned *M* Gene Encodes Resistance Specificity

Although the *M* and *L6* genes are very similar, they interact with different rust avirulence genes, *A-M* and *A-L6*, respectively, that map to unlinked locations in the fungal genome (Lawrence et al., 1981). The cloned *M* gene, when introduced into Hoshangabad, which is susceptible to all rust strains so far tested, confers rust resistance with the *M* specificity. Similarly, the cloned *L6* gene, when introduced into Hoshangabad, confers resistance with *L6* specificity (J.G. Ellis and G.J. Lawrence, unpublished data). Thus, specificity is a property of these genes. Therefore, one or more differences between amino acid sequences in their products may indicate potential specificity determinants. The two amino acid sequences are 78% identical, and minor sequence differences occur throughout the polypeptide. However, when aligned (Figure 3), two regions were detected where major differences occur. The first is in the N-terminal region of the polypeptide in which 37 contiguous amino acid differences were observed between the *M* and *L6* products. However, only minor differences occur in this region of the *L2*, *L6*, and *L10* alleles, which also have distinct resistance

specificities (J.G. Ellis, unpublished data). The second region of the two polypeptides that shows major differences occurs in the product of exon 3. It is clear that simple sequence comparisons are inadequate to identify specificity determinants. Domain-swapping experiments that make chimeric genes will help to identify critical regions of the genes that determine specificity differences.

Origin of New Specificities

It has been proposed (Pryor, 1987) but not demonstrated that new resistance specificities may arise at complex loci by unequal exchange, resulting in the reassortment of pre-existing genetic information. Recently, Richter et al. (1995) described the first experimental generation of new rust resistance specificities at the complex *Rp1* locus in maize and showed that these events were associated with exchange of flanking restriction fragment length polymorphism markers. Description of the actual process by which these new specificities arose will require the cloning of the genes involved.

In this article, we described three independent, spontaneous mutants in 270,000 progeny involving the two direct repeat units in the 3' end of the *M* gene coding region. The altered *M* alleles encode an internally deleted *M* protein containing a single hybrid repeat unit. These mutants do not express the *M* rust resistance specificity, which indicates that the two-repeat structure is essential for resistance gene function or specificity. Because the change in each independent mutant was identical, we hypothesize that either an intragenic recombination event between the two repeats or slipped alignment during replication (e.g., see Lovett and Feschenko, 1996) gave rise to the deletions. DNA sequence alignment of the two wild-type repeat units and the single repeat in the mutants (data not shown) limit the recombination exchange site or realignment site to within a 45-bp region that is invariant between the two repeats (Figure 4B). This exchange event was almost certainly intragenic and not between tandem, related genes in the complex *M* locus. The evidence for this is the sequence identity (data not shown) in the wild-type and *m-x39* alleles in the region 3' of the coding region, containing the *M* gene-specific and single-copy *Mxb-1* probe. In addition, in *Xba*I digests (Figure 1A), only a single *M*-related fragment was altered. This region was not sequenced in the *m-x32* and *m-x123* alleles; however, the presence of this sequence in the 3.0-kb *Sal*I fragments cloned from both mutants was confirmed by DNA hybridization by using probe *Mxb-1* (data not shown).

Initial examination of one of the mutant alleles, *m-x32*, with a rust strain that recognizes many different rust resistance genes did not detect a new specificity. Nevertheless, the result shows that the terminal direct repeat units may play an active role in the evolution of the structure of these genes.

Two potential mechanisms to generate these mutants have been considered: unequal crossing over at meiosis or

slipped alignment during replication. Although the three *M* mutants described in this work involved reduction in copy number of the repeat unit, both mechanisms can generate increases in copy number (Lovett and Feschenko, 1996). Such processes may account for the presence of four terminal repeat units that occur in the *L2* rust resistance allele (J.G. Ellis, unpublished data). Dixon et al. (1996) have also described a pronounced reiterated DNA sequence in the LLR-encoding resistance gene *Cf-2* and considered the effect of these repeats in leading to an increased rate of intragenic unequal exchange and the potential of such events to generate variation in resistance specificity.

METHODS

Plant Material

The flax line Forge is homozygous for four rust resistance genes *L6*, *M*, *N*, and *P2* (Lawrence et al., 1989). The line Hoshangabad contains no detectable rust resistance specificity (Mayo and Shepherd, 1980). The line Trac1, which is homozygous for a single *Activator (Ac)* element inserted in the *L6* gene, was obtained by crossing the *Ac*-tagged mutant X75 (Lawrence et al., 1995) to line Birio (*L6*) and selecting from among the resulting hybrids a plant that contained only the *Ac* in *L6*. Selfed progeny from this plant were screened with rust strain CH5F2-84, which is avirulent on *L6*. A susceptible plant that lacked *L6* and therefore was homozygous for the *Ac*-tagged *L6* allele was used to establish the Trac1 line.

The flax lines Cass (*M3*), Bison (*L9*), and Dakota (*M*) have been described by Islam and Mayo (1990). A line of flax homozygous for the mutant *M* allele *m-x32* but that lacked the *L6*, *N*, and *P2* resistance genes was derived from descendants of mutant plant X32 (this study). Progeny that were homozygous for *m-x32* and *LH*, an allele at the *L* locus in Hoshangabad that is undetected by any known strain of flax rust, were identified by restriction fragment length polymorphism analysis by using the markers Lu-2-1B and Lu-2-2A, which are linked to the *L* and *M* loci, respectively (Ellis et al., 1995). From among these progeny, a plant lacking the *P2* gene was identified by its susceptibility to rust strain CH5F2-133. Among the progeny of this plant, an individual lacking the *N* gene was identified by its susceptibility to rust strain P1C.

Rust Strains

The rust strains used in this study were described previously (Lawrence et al., 1981).

Transposon Tagging

The strategy for obtaining *Ac*-tagged mutants of a rust resistance gene in flax has been described elsewhere (Lawrence et al., 1993, 1995). It involved crossing Forge plants that carried *Ac* or *Tam3* elements to Hoshangabad and screening the progeny for individuals susceptible to a rust strain, CH5F2-78, that recognizes the *M* gene (but not the *L6*, *N*, and *P2* genes). Some of the tagging experiments involved derivatives of Forge lines that contained a high copy num-

ber of *Ac*, one of which was linked to *M*. The *Ac* linked to the *M* gene (1 map unit) was identified initially as the sole *Ac* in a primary transformant of Forge (E.J. Finnegan, J.G. Ellis, and G.J. Lawrence, unpublished results). Because *Ac* appears to be inactive in flax when present in one or a few copies (Ellis et al., 1992), this *M*-linked *Ac* was backcrossed into a line of flax (D97) possessing ~15 copies of *Ac* (Lawrence et al., 1993). The BC1, BC2, BC3, and BC3-selfed progeny were crossed to Hoshangabad to produce progeny for screening for mutants. Mutant X142, which contained an *Ac*-tagged *M* gene, was derived from one of the BC3-selfed individuals.

Protoplast Preparation and Pulsed-Field Gel Electrophoresis

Two grams of apical leaves was harvested from 30- to 40-day-old flax seedlings, cut into 1-mm strips with a razor blade, and incubated overnight with protoplasting solution (Zhan et al., 1989). Protoplasts were separated from undigested leaf material by passage through a 70- μ M mesh and harvested by centrifugation at 100g for 15 min at room temperature. The protoplasts were resuspended in 10 mL of SCE (1 M sorbitol, 60 mM sodium citrate, and 1 mM EDTA, pH 7.0), centrifuged again, and resuspended in an equal volume of SCE. Agarose microbeads were made from the SCE protoplast suspension, proteinase treated, and digested with restriction enzymes (Wing et al., 1993). The digested DNA was separated by pulsed-field gel electrophoresis in a TAFE apparatus (Beckman, Palo Alto, CA) for 12 hr at 350 mA with a 1-min pulse time followed by 6 hr at 370 mA with 2-min pulses. Intact chromosomes of *Saccharomyces cerevisiae* were used as size standards. After electrophoresis, we stained the gel with 0.5 μ g/mL ethidium bromide for 1 hr at room temperature, and the DNA was nicked by UV irradiation (600 mJ in a Stratilinker UV cross-linker; Stratagene, La Jolla, CA) before transfer. All DNA blots and hybridizations were done according to previously described methods (Taylor et al., 1989).

Transcript Analysis

The cDNA library of mRNA from Forge plants was described by Lawrence et al. (1995). Additional *M* gene-specific cDNA was obtained by long-range polymerase chain reaction (PCR), using a Perkin-Elmer kit and conditions recommended by the manufacturer, except that a higher Mg^{2+} concentration (4 mM) was used. Amplification conditions were 94°C for 5 min followed by 30 cycles of 94°C for 30 sec and 62°C for 6 min and then one cycle of 72°C for 10 min in 100- μ L reactions. The substrate for PCR was total cDNA. Primers were 5'-GGAATTCTTGCACGTTTCATCATTGAAGGAATG-3' that incorporated the ATG translation initiator and 5'-GGAATTCATCGAAAGTTAATGTACTACAGTAG-3' from downstream of the stop codon. The initial product of amplification was reamplified (1 μ L in 100 μ L), digested with XbaI and SacI that cut internally in the *M* genomic sequence upstream of intron 2 and downstream of intron 3, and then cloned into pUC119 for sequence analysis.

Genomic DNA Blot Analysis

Procedures for flax genomic DNA blot analysis were hybridization in $5 \times$ SSC (1 \times SSC is 0.15 M NaCl, 0.015 M sodium citrate) at 42°C

and in 50% formamide with final filter washes in $0.1 \times$ SSC at room temperature, as described previously (Ellis et al., 1992). Probes critical to this analysis were Lu-3 (nucleotide coordinates 2917 [EcoRI] to 3916 [XbaI] from the leucine-rich repeat [LRR] region of *L6* [Lawrence et al., 1995]), a probe derived from the 3' end of *Ac* (Ellis et al., 1992), and Mxb-1, a 700-bp XbaI fragment derived from the region 3' of the *M* gene (this study).

Genomic Cloning

Flax DNA (extracted according to methods described in the Bethesda Research Laboratories Focus publication, Vol. 4, December 1982) was cut with either Sall or EcoRI. The DNA was size fractionated on sucrose gradients (10 to 40%) in 13-mL polyallomer tubes and centrifuged at 30,000 rpm at 20°C for 22 hr in an Sw-41 rotor (Beckman). Samples of 0.5 mL were collected from the gradients, and Sall-digested DNA from the appropriate size range was ligated to bacteriophage EMBL4 DNA digested with Sall. A similar ligation was set up for EcoRI-digested DNA from mutant X142, with EMBL3 DNA being cut with both EcoRI and BamHI. After packaging and plating on *Escherichia coli* (strain K803), positive plaques from the Sall fragment library were identified by hybridization with the *L6* DNA probe, Lu-3 (Ellis et al., 1995), or in the case of the *Ac*-tagged X142 mutant, a 3' end probe from *Ac* (Ellis et al., 1992). The flax genomic library of partially digested Sau3A DNA from which the wild-type *M* gene was cloned using probe Mxb-1 (see Figure 2A, map ii) is described elsewhere (Lawrence et al., 1995).

DNA Sequencing and Computer Analysis

All DNA sequence data were obtained from PCR cycle sequencing using an ABI dye primer sequencing kit with overlapping sequential deletions made by exonuclease III (Pharmacia). Nucleotide and amino acid sequence data were compiled and analyzed using the Genetics Computer Group (Madison, WI) sequence analysis programs (Devereux et al., 1984).

Flax Transformation

Flax was transformed essentially as described by Lawrence et al. (1989), except that spectinomycin sulfate (100 μ g/mL) was used for selection. An EcoRI fragment containing the entire *M* gene was cloned into the binary vector pTAB-EPspec (J.G. Ellis, unpublished data) between the right T-DNA border and a plant-selectable spectinomycin resistance gene inserted near the left T-DNA border.

ACKNOWLEDGMENTS

Excellent technical assistance was provided by Valerie Ryle, Margaret Keys, and Patricia Atkinson. The support and critical advice of A.J. Pryor is gratefully acknowledged. This research was supported by grants from the Grains Research and Development Corporation.

Received November 27, 1996; accepted January 30, 1997.

REFERENCES

- Bent, A.F., Kunkel, B.N., Dahlbeck, D., Brown, K.L., Schmidt, R., Giraudat, J., Leung, J., and Staskawicz, B.J. (1994). *RPS2* of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. *Science* **265**, 1856–1860.
- Boyes, D., McDowell, J., and Dangl, J. (1996). Many roads lead to resistance. *Curr. Biol.* **6**, 634–637.
- Devereux, J., Haeblerli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**, 387–395.
- Dixon, M.S., Jones, D.A., Keddle, 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.
- Ellis, J.G., Lawrence, G.J., and Finnegan, E.J. (1992). Developing a transposon tagging system to isolate rust resistance genes from flax. *Theor. Appl. Genet.* **85**, 46–54.
- Ellis, J.G., Lawrence, G.J., Finnegan, E.J., and Anderson, P.A. (1995). Contrasting complexity of two rust resistance loci in flax. *Proc. Natl. Acad. Sci. USA* **92**, 4185–4188.
- Flor, H.H. (1954). Seed-flax improvement. III. Flax rust. *Adv. Agron.* **6**, 152–161.
- Flor, H.H. (1965). Tests for allelism of rust resistance genes in flax. *Crop Sci.* **5**, 415–418.
- Grant, M.R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.W., and Dangl, J.L. (1995). Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. *Science* **269**, 843–846.
- Islam, M.R., and Mayo, G.M.E. (1990). A compendium on host genes in flax conferring resistance to flax rust. *Plant Breed.* **104**, 89–100.
- Islam, M.R., and Shepherd, K.W. (1991). Present status of genetics of rust resistance in flax. *Euphytica* **55**, 255–267.
- Islam, M.R., Shepherd, K.W., and Mayo, G.M.E. (1989). Recombination among genes at the *L* group in flax conferring resistance to rust. *Theor. Appl. Genet.* **77**, 540–546.
- Jones, D.A., and Jones, J.D.G. (1997). The roles of leucine-rich repeat proteins in plant defences. *Adv. Bot. Res. Adv. Plant Pathol.* **24**, 89–167.
- Jones, D.A., Thomas, C.M., Hammond-Kosack, K.E., Balint-Kurti, P.J., and Jones, J.D.G. (1994). Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging. *Science* **266**, 789–793.
- Kobe, B., and Deisenhoffer, J. (1994). The leucine-rich repeat: A versatile binding motif. *Trends Biochem. Sci.* **19**, 415–421.
- Lawrence, G.J., Mayo, G.M.E., and Shepherd, K.W. (1981). Interactions between genes controlling pathogenicity in the flax rust fungus. *Phytopathology* **71**, 12–19.
- Lawrence, G.J., Ellis, J.G., Finnegan, E.J., Dennis, E.S., and Peacock, W.J. (1989). Tagging rust resistance genes in flax. In *Breeding Research: The Key to the Survival of the Earth*, S. Iyama and G. Takeda, eds (Tsukuba, Japan: Organizing Committee of the Sixth International Congress SABRAO), pp. 535–538.

- Lawrence, G.J., Finnegan, E.J., and Ellis, J.G.** (1993). Instability of the *L6* gene for rust resistance is correlated with the presence of a linked *Ac* element. *Plant J.* **4**, 659–669.
- Lawrence, G.J., Finnegan, E.J., Ayliffe, M.A., and Ellis, J.G.** (1995). The *L6* gene for flax rust resistance is related to the Arabidopsis bacterial resistance gene *RPS2* and the tobacco viral resistance gene *N*. *Plant Cell* **7**, 1195–1206.
- Lovett, S.T., and Feschenko, V.V.** (1996). Stabilization of diverged tandem repeats by mismatch repair: Evidence for deletion formation via a misaligned replication intermediate. *Proc. Natl. Acad. Sci. USA* **93**, 7120–7124.
- Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganai, M.W., Spivey, R., Wu, T., Earle, E.D., and Tanksley, S.** (1993). Map-based cloning of a protein kinase gene conferring disease resistance in tomato. *Science* **262**, 1432–1436.
- Martin, G.B., Frary, A., Wu, T., Brommonschenkel, S., Chunwongse, J., Earle, E.D., and Tanksley, S.** (1994). A member of the tomato *Pto* family confers sensitivity to fenthion resulting in rapid cell death. *Plant Cell* **6**, 1543–1552.
- Mayo, G.M.E., and Shepherd, K.W.** (1980). Studies of genes controlling specific host–pathogen interactions in flax and its rust. I. Fine structure analysis of the *M* group in the host. *Heredity* **44**, 212–227.
- Mindrinis, M., Katagiri, F., Yu, G.-L., and Ausubel, F.M.** (1994). The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* **78**, 1089–1099.
- Pryor, A.J.** (1987). The origin and structure of fungal disease resistance genes in plants. *Trends Genet.* **3**, 157–161.
- Pryor, A.J., and Ellis, J.G.** (1993). The genetic complexity of fungal disease resistance genes in plants. *Adv. Plant Pathol.* **10**, 281–305.
- Richter, T.E., Pryor, A.J., Bennetzen, J.L., and Hulbert, S.H.** (1995). New rust resistance specificities associated with recombination at the *Rp1* complex in maize. *Genetics* **141**, 373–381.
- 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 disease resistance genes and lies embedded within the *Pto* kinase gene cluster. *Cell* **86**, 123–133.
- Shepherd, K.W., and Mayo, G.M.E.** (1972). Genes conferring specific plant disease resistance. *Science* **175**, 375–380.
- Staskawicz, B.J., Ausubel, F.M., Baker, B.J., Ellis, J.G., and Jones, J.D.G.** (1995). Molecular genetics of plant disease resistance. *Science* **268**, 661–667.
- Taylor, B.H., Finnegan, E.J., Dennis, E.S., and Peacock, W.J.** (1989). The maize transposable element *Ac* excises in progeny of transformed tobacco. *Plant Mol. Biol.* **13**, 109–118.
- Whitham, S., Dinesh-Kumar, S.P., Choi, D., Hehl, R., Corr, C., and Baker, B.** (1994). The product of the tobacco mosaic virus gene *N*: Similarity to Toll and interleukin-1 receptor. *Cell* **78**, 1101–1115.
- Wing, R.A., Rastogi, V.K., Zhang, H.B., Paterson, A.H., and Tanksley, S.D.** (1993). An improved method of plant megabase DNA isolation in agarose microbeads suitable for physical mapping and YAC cloning. *Plant J.* **4**, 893–898.
- Zhan, X.-C., Jones, D.A., and Kerr, A.** (1989). *In vitro* plantlet formation in *Linum marginale* from cotyledons, hypocotyls, leaves, roots and protoplasts. *Aust. J. Plant Physiol.* **16**, 315–320.