

# The Tomato *Cf-2* Disease Resistance Locus Comprises Two Functional Genes Encoding Leucine-Rich Repeat Proteins

Mark S. Dixon, David A. Jones, James S. Keddle,\*  
Colwyn M. Thomas, Kate Harrison,  
and Jonathan D. G. Jones  
Sainsbury Laboratory  
John Innes Centre  
Colney Lane  
Norwich, NR4 7UH  
United Kingdom

## Summary

In plants, resistance to pathogens is frequently determined by dominant resistance genes, whose products are proposed to recognize pathogen-encoded avirulence gene (*Avr*) products. The tomato resistance locus *Cf-2* was isolated by positional cloning and found to contain two almost identical genes, each conferring resistance to isolates of tomato leaf mould (*C. fulvum*) expressing the corresponding *Avr2* gene. The two *Cf-2* genes encode protein products that differ from each other by only three amino acids and contain 38 leucine-rich repeat (LRR) motifs. Of the LRRs, 20 show extremely conserved alternating repeats. The C-terminus of *Cf-2* carries regions of pronounced homology to the protein encoded by the unlinked *Cf-9* gene. We suggest that this conserved region interacts with other proteins involved in activating plant defense mechanisms.

## Introduction

Resistance (*R*) genes in plants play a crucial role in preventing disease. Many *R* genes are dominant, or incompletely dominant, and require specific dominant avirulence (*Avr*) genes in the pathogen for their function (Flor, 1946). This genetic interaction between plant and pathogen led to the current view that such *R* genes encode receptors for *Avr* gene-dependent pathogen molecules (reviewed by Staskawicz et al., 1995). Upon recognition of these molecules, *R* gene products activate plant defense mechanisms. These defenses include rapid production of an oxidative burst resulting in cell wall cross-linking, localized cell death (the hypersensitive response), salicylic acid biosynthesis, and induction of genes characteristic of systemic acquired resistance (Levine et al., 1994; Ward et al., 1991; Lamb, 1994).

*R* genes fall into several distinct classes. The tobacco *N* gene, the Arabidopsis *RPS2* and *RPM1* genes, and the flax *L6* gene confer resistance, respectively, to viral, bacterial, and fungal pathogens, but they all encode proteins that are probably cytoplasmic, contain multiple leucine-rich repeats (LRRs), and a nucleotide-binding site (Whitham et al., 1994; Bent et al., 1994; Mindrinos et al., 1994; Grant et al., 1995; Lawrence et al., 1995).

The tomato *Cf-9* gene confers resistance to isolates of *Cladosporium fulvum*, which express the corresponding *Avr9* gene. This has been the sole member of a second class, predicted to be predominantly extracytoplasmic, with a C-terminal membrane anchor and no nucleotide-binding site (Jones et al., 1994). The extracytoplasmic domain consists primarily of LRRs, and the predicted cytoplasmic domain of *Cf-9* is only 25 amino acids. The tomato *Pto* gene, conferring resistance to a bacterial pathogen, is a member of a third class and encodes a serine/threonine protein kinase, but lacks LRRs (Martin et al., 1993). The rice *Xa21* gene, also conferring resistance to a bacterial pathogen, carries features of both the latter classes, encoding a transmembrane protein kinase with 23 extracellular LRRs (Song et al., 1995).

To understand how *R* gene products recognize specific pathogen molecules and activate plant defenses, we have studied the interaction between cultivated tomato (*Lycopersicon esculentum*) and *C. fulvum*. Several different *R* genes that confer resistance to different races of *C. fulvum* have been reported (Stevens and Rick, 1988), of which *Cf-2*, *Cf-3*, *Cf-4*, *Cf-5*, and *Cf-9* have received particular attention (Jones et al., 1993; Hammond-Kosack and Jones, 1994), leading to the isolation of *Cf-9* (Jones et al., 1994). Several *Cf* genes have been identified in wild species and subsequently bred into cultivated tomato to generate near isogenic lines (NILs; see Experimental Procedures). Both *Cf-2* and *Cf-9* were identified in *Lycopersicon pimpinellifolium* species, whereas *Cf-5* was identified in the land race, *L. esculentum* var. *cerasiforme*. Two *Avr* genes, *Avr4* and *Avr9*, have been cloned from *C. fulvum* and shown to encode small, secreted cysteine-rich peptides, which are candidate ligands for the *Cf-4* and *Cf-9* gene products, respectively (Joosten et al., 1994; van Kan et al., 1991). *C. fulvum* races carrying *Avr4* or *Avr9* are avirulent on tomato lines that carry the corresponding *R* genes *Cf-4* or *Cf-9*.

To understand *Cf* gene function, it is clearly desirable to compare the sequences of several different *R* gene products. Regions that have high homology between different *R* genes might play a role in conserved functions such as initiating a signal transduction chain, whereas divergent regions might participate in specific binding to a unique ligand. After isolating *Cf-9*, our objective has been to isolate other *Cf* genes to allow such comparisons to be made.

## Results

### A Rare Recombinant between *Cf-2* and *Cf-5* Defines a Physical Location for *Cf-2*

Earlier work mapped the *Cf-2* resistance locus to chromosome 6 between the restriction fragment length polymorphism (RFLP) marker GP79 and the isozyme marker *Aps-1* (Dickinson et al., 1993). Recently, *Cf-2* has been more precisely positioned between the RFLP loci MG112 and CT119, which in *Cf0* *L. esculentum* (susceptible to all races of *C. fulvum*) are only 30 kb apart (Dixon

\*Present address: Department of Plant Biology, University of California, Berkeley, Berkeley, California 94720.

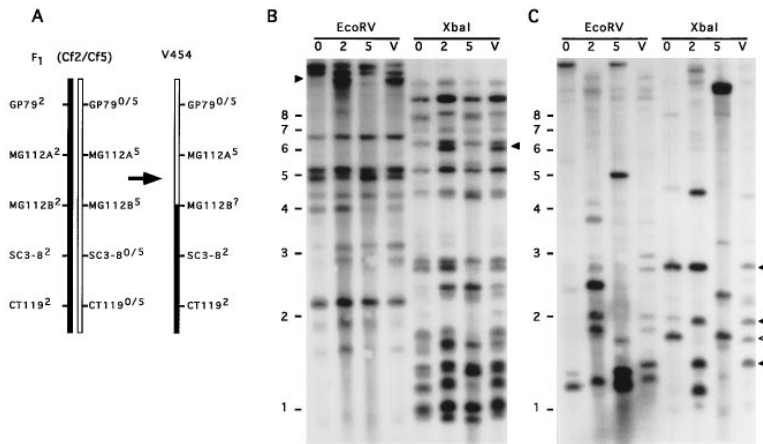


Figure 1. Analysis of a Recombinant between *Cf-2* and *Cf-5*

(A) Schematic representations of regions of chromosome 6 are shown for key plants in the crosses used to identify the rare recombinant between *Cf-2* and *Cf-5* (V454). The linear order of closely linked RFLP probes are shown, and the genotype of each marker is denoted by superscripted text. The 0/5 superscript indicates that no polymorphism has been identified that reveals a difference between the Cf0 and Cf5 NILs; the question mark indicates that the marker reveals a pattern that does not conform entirely to that of any of the parental genotypes. F1 designates the plant derived from crossing Cf2 to Cf5, while V454 indicates the recombinant chromosome identified in the single disease-sensitive plant (V454) among 12,000 progeny of a test cross of the F1 heterozygote to Cf0.

(B) Hybridization of the RFLP marker SC3-8 to Cf0 (0), Cf2 (2), Cf5 (5), and a homozygote for the V454 recombinant chromosome (V). Closed arrowheads indicate the positions of bands specific for *L. pimpinellifolium* DNA introgressed with *Cf-2*. Molecular weights in kilobases are indicated to the left.

(C) The same blot as in (B) hybridized with a 5.5 kb HindIII probe derived from Cf0 encompassing part of the region designated as MG112B together with some flanking sequences. Closed arrowheads identify Cf2-specific bands that are present in the recombinant V454, and the open arrowhead identifies a novel band not present in either the Cf2 or Cf5 parents.

et al., 1995). The probe MG112 hybridizes strongly to a location in *L. esculentum* (Cf0) DNA designated MG112A and more weakly to an adjacent (~5 kb proximal) location MG112B. Analysis of DNA from plants recombinant for RFLP markers around this region place *Cf-2* between the markers MG112A and CT119. A cosmid contig spanning the equivalent position in the susceptible Cf0 genome was established and carries regions corresponding to the locus MG112B. This cloned region in Cf0, which includes the region designated MG112B, was hypothesized to carry sequences homologous to *Cf-2* (Dixon et al., 1995). The order of RFLP markers within this region is indicated in Figure 1A. The *Cf-5* gene, derived from *L. esculentum* var. *cerasiforme*, has been assigned to the same map location and is either allelic with, or very closely linked to, *Cf-2* (Jones et al., 1993; Dickinson et al., 1993).

To investigate further the close linkage of *Cf-2* and *Cf-5*, and to exploit this proximity for high resolution analysis by recombination, we crossed a line containing *Cf-2* (Cf2) to a line containing *Cf-5* (Cf5) to generate a doubly heterozygous F1 plant carrying a single copy of each resistance locus (Figure 1A). This F1 was test crossed to a line lacking any detectable *Cf* genes (Cf0), and approximately 12,000 test cross-progeny were screened for susceptible plants by inoculation with a race of *C. fulvum* (race 4 GUS) that expresses the *Escherichia coli*  $\beta$ -glucuronidase (GUS) gene (Oliver et al., 1993). All progeny should carry either *Cf-2* or *Cf-5* and be resistant except where recombination has occurred between the two. A single susceptible individual (V454) was identified and shown to carry a recombinant chromosome 6 by using RFLP markers known to be closely linked to *Cf-2* (Dixon et al., 1995) (Figure 1A). To simplify subsequent RFLP analysis, we identified an F2 individual homozygous for the recombinant chromosome from self-progeny of V454. The RFLP marker GP79, which maps distal to *Cf-2*, reveals an *L. esculentum* hybridization pattern in V454 (data not shown). The probe MG112 identifies polymorphisms between the NILs Cf0, Cf2, and Cf5 with all restriction enzymes tested (Dixon et al.,

1995) and confirms the MG112A region as *L. esculentum* var. *cerasiforme* in V454 (data not shown) and thus distal to *Cf-2*. Two polymorphic RFLP markers, SC3-8 (Figure 1B) and CT119, which map nearby on the proximal side of *Cf-2*, are both genotypically *L. pimpinellifolium* (Cf2) in V454, indicating that recombination has occurred between SC3-8 and MG112A. SC3-8 identifies a small family of related sequences, one of which shows a polymorphism between the Cf0 and Cf2 NILs (Figure 1B). These data rule out MG112A as a candidate for *Cf-5* and SC3-8 as a candidate for *Cf-2*, since the V454 plant that carries these two alleles is susceptible to races of *C. fulvum* expressing the corresponding *Avr2* and *Avr5* genes.

A probe (probe C) comprising a 5.5 kb HindIII fragment derived from Cf0, which contains part of the region designated as MG112B together with some flanking sequences (Figure 2A), reveals a hybridization pattern in V454 different from both the parental Cf2 and Cf5 genotypes (Figure 1C). For example, XbaI-digested V454 DNA contains many of the probe C-homologous bands characteristic of Cf2 (Figure 1C, closed arrowheads). However, one band is missing, and a new band, absent from either parental genotype, is apparent (Figure 1C, open arrowhead). This suggests that the recombination event in V454 took place in, or close to, a region of probe C homology and implies that probe C could hybridize to both the *Cf-2* and *Cf-5* genes. The same blot probed with SC3-8 indicates approximately equal DNA loading compared with Cf2 and Cf5 (Figure 1B). The reduced hybridization to probe C in the V454 recombinant (Figure 1C) could have arisen if *Cf-2* and *Cf-5* are both part of a multigene family and V454 arose by an unequal crossover event that reduced the copy number in the recombinant chromosome to less than that in either parent.

#### Isolation of Binary Vector Cosmid Clones Carrying *Cf-2*

To clone *Cf-2*, we used genomic DNA from plants homozygous for the *Cf-2* and *Cf-9* resistance genes to construct a binary vector cosmid library. We employed the

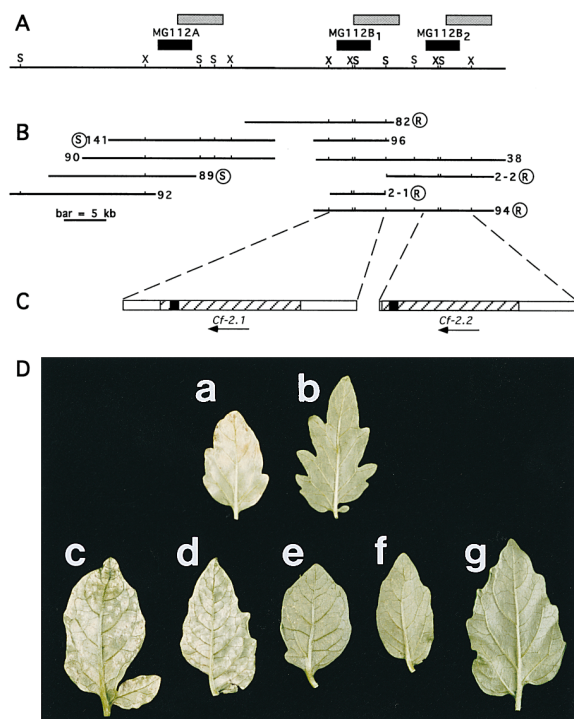


Figure 2. Complementation with Cosmids from the *Cf-2* Locus  
(A) Map of *Sacl* (S) and *XhoI* (X) restriction endonuclease sites in the region carrying overlapping cosmids that hybridize to MG112. Closed boxes indicate the regions (A and B) that hybridize with the RFLP marker MG112, while stippled boxes indicate the regions that hybridize with the 5.5 kb *HindIII* probe derived from *Cf0* (probe C) used for Southern blot analysis in Figure 1C.  
(B) Cosmid contig spanning the *Cf-2* locus. Horizontal lines represent the inserts contained within eight independent MG112-hybridizing cosmid clones. The circled letters R and S identify cosmids or subclones that either confer resistance or fail to confer resistance (susceptible), respectively, upon transformation. The scale bar indicates 5 kb.  
(C) Sequenced regions of cosmid 94 are shown as rectangles. Hatching indicates the transcribed regions, and closed boxes indicate introns. Horizontal arrows show the direction of transcription. Broken lines indicate the location of the sequenced regions with respect to cosmid 94.  
(D) Leaves of wild-type plants and plants transformed with cosmid 94 after challenge with *C. fulvum* race 4. Shown are *Cf0* (a), *Cf2* (b), and the first five independent transformants with cosmid 94 (c-g). Southern blot analysis indicates that transformants shown in (c) and (d) do not contain an intact copy of *Cf-2*.

binary cosmid vector pCLD04541 (Bent et al., 1994), which carries a kanamycin resistance selectable marker and a multiple restriction enzyme cloning site between the left and right T-DNA borders. These features of pCLD04541 allow the direct *Agrobacterium*-mediated transfer of cloned plant DNA into plant cells for complementation experiments.

Upon screening of the library with RFLP probe MG112, eight strongly hybridizing pools were identified (see Experimental Procedures). Clones were plated from each of the corresponding glycerol stocks, and individual clones that hybridized strongly were isolated by standard procedures. By a combination of restriction mapping, fingerprinting, and DNA hybridization, the cloned insert DNA in each cosmid was mapped and assembled

into a single contig (Figure 2A). This map shows colinearity with the corresponding region from *L. esculentum* (Dixon et al., 1995). However, there are two MG112B regions with similar restriction maps, which could represent a gene duplication in *L. pimpinellifolium* relative to *Cf0* (Figure 2A). All major MG112- and probe C-hybridizing bands on genomic Southern blots can be explained by the sequences contained within this *Cf-2* cosmid contig.

#### Complementation in Transgenic Plants Reveals Two Functional *Cf-2* Genes

Four independent clones (cosmids 82, 89, 94, and 141) were mobilized into *Agrobacterium* and used to transform susceptible tomato plants. Transgenic plants were selected and screened for resistance to race 4 GUS *C. fulvum* (Figure 2D). All four plants carrying cosmid 82, and 11 of 16 plants transformed with cosmid 94, were resistant to *C. fulvum* (Figure 2D). Fungal biomass was estimated by staining for GUS activity using X-Gluc (5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronide cycloammonium salt) and showed that, at a macroscopic level, resistant transgenic plants were as resistant as control *Cf2* plants. All transgenic plants carrying cosmids 89 and 141 remained fully susceptible. These data define the region containing the *Cf-2* gene as the overlap between cosmids 82 and 94. This region is physically defined as MG112B<sub>1</sub> (Figure 2A).

Upon transfer of cosmid 94 from the recombination-deficient cloning host *E. coli* SURE tet<sup>r</sup> (Stratagene) to *Agrobacterium*, the insert DNA consistently underwent recombination between the two copies of the MG112B region, such that a chimeric copy of MG112B was always created (data not shown). Despite this, the resulting transgenic plants were still resistant to *C. fulvum*, suggesting the possibility that both copies of MG112B may in fact be functional *Cf-2* genes.

Each copy of MG112B was subcloned to generate two new binary vector clones, 2-1 carrying a minimal MG112B<sub>1</sub> region and 2-2 carrying MG112B<sub>2</sub> (Figure 2B). Transformation of sensitive plants with these clones gave rise to five resistant transformants out of five with clone 2-1, and 16 resistant transformants out of 18 with clone 2-2. These data prove the existence of two independently functional *Cf-2* resistance genes.

Progeny from all resistant nonpolyploid transformed plants were screened with matched races of *C. fulvum* either containing or lacking *Avr2* (race 5,9 compared with race 2,5,9). Races of *C. fulvum* are named after the resistance genes they can overcome. All progeny were susceptible to *C. fulvum* lacking *Avr2* (race 2,5,9), whereas approximately 75% of progeny from each transformant were resistant to *C. fulvum* carrying *Avr2* (race 5,9) (data not shown). These data confirm the race-specific nature of the resistance genes cloned as *Cf-2*.

#### Structure and Expression of the *Cf-2* Genes

The DNA sequences of both *Cf-2* genes, indicated by the boxed regions in Figure 2C, were determined. Each copy contains a single major open reading frame encoding a polypeptide of 1112 amino acids. Only three nucleotides differ between the open reading frames, each

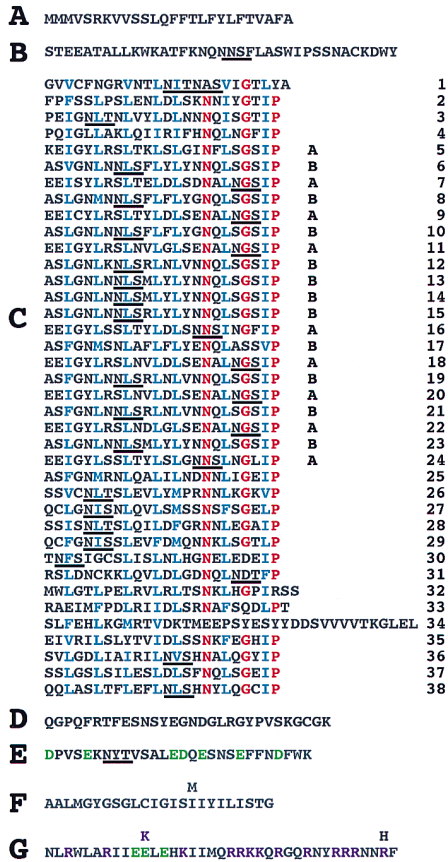


Figure 3. Primary Structure of the Cf-2 Proteins  
The amino acid sequences predicted from the DNA sequence of Cf-2.2 are shown divided into seven domains (A–G), as described in the text. In domain C, the conserved L of the LRRs is often replaced by V, F, I, or M. These are highlighted in blue, while other conserved amino acids are shown in red. In domains E and G, acidic amino acids are highlighted in green, and the basic amino acids are in purple. Potential N-linked glycosylation sites are underlined. The three amino acid differences in Cf-2.1 are placed above their corresponding position in the sequence of Cf-2.2. Numbers to the right of domain C indicate the specific LRR number. Letters A and B to the right of the sequence indicate the type of LRR (see text).

giving rise to amino acid changes (Figure 3). The sequences are identical for 787 bases upstream of the first of three consecutive methionine codons and beyond that show 95.2% similarity until the end of the determined sequence (~0.9 kb). The genomic DNA sequence 3' to the termination codon to the poly(A) addition site (~0.4 kb) shows 95.3% identity.

When compared with sequences in the GenBank database, Cf-2 shows greatest homology (BlastX score of 483; probability of  $1.3 \times e^{-152}$ ) (Altschul et al., 1990) to Cf-9 (Jones et al., 1994) and can be subdivided into similar functional domains (Figure 3). The major portion of Cf-2, domains B, C, and D, has homology to the polygalacturonase inhibitor proteins (PGIPs) (Jones et al., 1994), with domain C containing 33 perfect and five imperfect LRRs. There are 31 potential N-linked glycosylation sites within domains B, C, and E. Domain A is a putative signal peptide of 26 amino acids. The predicted

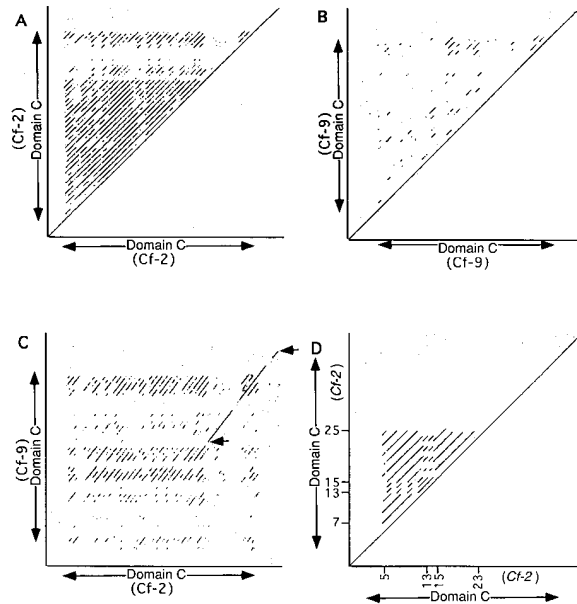


Figure 4. Analysis of Repeated and Homologous Sequences within Cf-2 and Cf-9

(A–C) The Cf-2.1 and Cf-9 protein sequences were compared with themselves and each other using the Genetics Computer Group program COMPARE (University of Wisconsin). A window of 24 and a stringency of 18 (75% similarity) was used for each analysis, and results are displayed as diagan plots. The limits of domain C (the LRR region) in each protein sequence are indicated on each axis. The N-terminus of each protein is aligned with the origin of each graph.

(A) The Cf-2.1 protein compared with itself.  
(B) The Cf-9 protein compared with itself.  
(C) Comparison of Cf-2.1 protein with Cf-9 protein. Arrows identify the extent of the high homology in the C-terminal region (see text).  
(D) The Cf-2.1 nucleic acid sequence encompassing the entire open reading frame was compared with itself. A window of 72 (one complete LRR) and a stringency of 54 (75% identity) was chosen. The positions for the starts of particular LRRs are indicated on each axis.

signal peptide cleavage site is based on the rules described by von Heijne (1986) and upon the beginning of the homology with the PGIPs. Domain F is a putative membrane-spanning region of 24 noncharged amino acids. Domains E and G are rich in acidic and basic residues, respectively, consistent with the role of anchoring and orienting the transmembrane domain within the cell membrane. Analysis of the Cf-2 protein using COMPARE (Genetics Computer Group, University of Wisconsin) reveals a highly reiterated repeat structure through domain C (Figure 4A), which is much more apparent than in Cf-9 (Figure 4B). This internal repeat structure is particularly conspicuous when diagan plot analysis is conducted on the DNA sequence (Figure 4D). Diagan plot analysis comparing Cf-2 with Cf-9 (Figure 4C) shows two diagonal lines in the region including the last 9.5 LRRs to the C-termini of both proteins, indicating strong homology. BESTFIT analysis (Genetics Computer Group) identifies these homologies in greater detail and indicates that they extend from the third asparagine in LRR 29 of Cf-2 (Figure 5). Within this region of similarity, a few specific regions of identity occur. An eight amino



in a necrotic phenotype in resistant transgenics. However, no necrosis matching that of *ne* was observed in any primary transformants, indicating that the necrotic phenotype is not dependent upon the presence of *Cf-2* alone, but on another linked dominant gene carried on the same introgressed region. This is consistent with the observation that the line Ontario 7620, which carries *Cf-2* on a small introgressed segment, does not display the necrotic phenotype (D. A. J., unpublished data).

### The LRR Domain of *Cf-2*

The *Cf-2* gene encodes the longest uninterrupted stretch of LRRs so far observed (Kobe and Deisenhofer, 1994). A particularly striking feature of the *Cf-2* proteins is the regularity of the length of the LRRs and the strong conservation of certain amino acids within the LRRs. Several of the LRRs can be grouped into two classes, A and B, based upon their sequence; the A class has the consensus EEIXYLXSLXXLXSENALNGSIP, while the B class has the consensus ASLGNLNNLSXLXLYXNQLSG SIP. Interestingly, these two classes of repeat often alternate with each other to give a second level of repeated structure (Figure 3). The two classes of LRR in *Cf-2* could act like the alternating A-type and B-type LRRs of porcine ribonuclease inhibitor protein (PRI) in providing interrepeat interactions that stabilize the tertiary structure (Kobe and Deisenhofer, 1993).

In PRI, neighboring repeats are rotated by 17° around an axis to produce a "horseshoe" tertiary structure (Kobe and Deisenhofer, 1993). In a protein with 21 or more copies of the PRI LRRs, the ends of the protein would collide. For *Cf-2*, with 38 repeats, a twist would be required to generate a structure similar to part of a coil. The tertiary structure of *Cf-2* might also differ from that of PRI for several other reasons. First, the PRI LRRs consist of alternating 28 and 29 amino acid repeats, which form a right-handed combination of a short  $\beta$  strand in the LXXLX region alternating with an  $\alpha$ -helical region whose core consists of three turns. The greater steric radius of the helix serves to ratchet each LRR by 17° (Kobe and Deisenhofer, 1993). In contrast, the LRRs of *Cf-2* and *Cf-9* are on average 24 amino acids long, and the potential helical region is consequently shorter. As in *Cf-9*, the LRRs of *Cf-2* contain a proline residue in a position of the repeat corresponding to the potential helical region. However, prolines can be readily incorporated near the N-termini of helices. More importantly, the spacing of leucine residues in this region differs considerably from PRI, suggesting that folding to form an amphipathic  $\alpha$  helix aligned parallel to the  $\beta$  strands as found in PRI is less likely. This is to be expected, since formation of a super-twisted repeated  $\beta$ - $\alpha$  structure, if it occurred, would necessitate some alteration of the basic helical region, perhaps both in length and relative orientation. Finally, the LRRs of *Cf-2* are predicted to be extracellular and have the potential to be highly glycosylated, unlike the LRRs of the cytoplasmically located PRI, and this could affect their tertiary structure. Kobe and Deisenhofer (1994) have speculated that some LRR proteins assume a  $\beta$ -helical structure analogous to that adopted by pectate lyase C; if this were true for *Cf-2* and *Cf-9*, they would adopt an alternative structure,

that of long rods. Preliminary molecular modeling studies (O. Kolade and A. Hemmings, personal communication) suggest that the short LRR motif of *Cf-2* and *Cf-9* cannot adopt a fold directly analogous to that of PRI. We cannot at present distinguish between the two possible alternative folds, but if *Cf-2* were to fold in the  $\beta$ -helical fashion over all 38 repeats, it would form a molecular rod approximately 185 Å long. Whichever structure *Cf-2* adopts, the LRR region could provide a structure on which a presentational surface can be supported and as well a mechanism by which diversity and novelty in the presentational surface can rapidly evolve.

### Comparison of the *Cf-2* and *Cf-9* Resistance Genes and Models for *Cf* Gene Function

The *Cf-2* and *Cf-9* proteins both confer recognition of *C. fulvum* Avr gene products, leading to initiation of the defense response. Activation by these two resistance genes leads to similar physiological changes: the production of an oxidative burst, lipid peroxidation, salicylic acid accumulation, and expression of several pathogenesis-related proteins (Vera-Estrella et al., 1992; Ashfield et al., 1994; May et al., 1996; Hammond-Kosack et al., 1994). A burst of superoxide produced by a mechanism analogous to the mammalian NADPH oxidase has been proposed to be an early event in many plant defense responses (Lamb, 1994). The production of an oxidative burst precedes the induction of salicylic acid by several hours in the tomato-*C. fulvum* interaction (May et al., 1996). Upon Avr recognition, *Cf* gene products may also activate a signal transduction pathway leading to the transcription of genes which contribute to plant defense mechanisms. Activation of an oxidative burst by the *Cf-2* and *Cf-9* gene products might occur through common signaling intermediates. Comparison of the *Cf-2* and *Cf-9* proteins reveals remarkable homology in their C-terminal portions and particularly in the last 9.5 LRRs of domain C (Figure 4C). It is striking that much of the conserved region between *Cf-2* and *Cf-9* is extracytoplasmic. This suggests two kinds of models for how this class of R gene might act (Figure 6). Transmission of a signal may involve Avr product-dependent interaction between this region of high identity and domains of a common or related membrane protein. One such protein might be the NADPH oxidase itself (Figure 6A). Another (Figure 6B) could be an LRR transmembrane receptor protein kinase, of which there are several examples in plants, such as TMK1, RLK5 (Chang et al., 1992; Walker, 1994), and the product of *Xa21*, the rice bacterial blight resistance gene (Song et al., 1995). Conceivably, in the same way that *Prf* and *Pto* are both required for *Pseudomonas* resistance in tomato (Salmeron et al., 1994; Martin et al., 1993), *Cf* function in tomato might require an *Xa21*-like gene, and *Xa21* function in rice might require a *Cf-9*-like gene. This model is reminiscent of *Brasica* self-incompatibility, which requires both a secreted stigmatic S locus glycoprotein and a stigmatic S-related kinase, with homology to S locus glycoprotein in its extracellular domain, for recognition of incompatible pollen (Nasrallah and Nasrallah, 1993).

Tests of these ideas about *Cf* gene function will

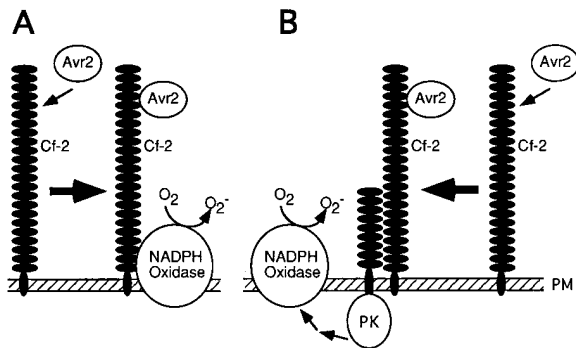


Figure 6. Two Models for Mechanisms by Which *Cf-2* and *Cf-9* Might Activate the Defense Response upon *Avr* Product Recognition PK, transmembrane protein kinase; PM, plasma membrane. In model (A), upon *Avr* recognition by a *Cf* gene product, the resulting complex might directly interact with a membrane-bound NADPH oxidase, which produces superoxide anions and initiates the plant defense response. In model (B), upon *Avr* product binding, the resulting complex interacts with a transmembrane protein kinase analogous to TMK1 or RLK5 (Chang et al., 1992; Walker, 1994) or to Xa21 (Song et al., 1995). Kinase action then triggers subsequent events that activate an NADPH oxidase.

emerge from future work. Genes required for *Cf-9* gene function (*rcr*, required for Cladosporium resistance) have been identified by mutagenesis (Hammond-Kosack et al., 1994). Currently, mutagenesis of *Cf2* plants is also being performed to identify other genes required for this resistance pathway; this general approach should identify partners that interact with the *Cf* gene products.

If the membrane-proximal conserved extracytoplasmic domains of *Cf-2* and *Cf-9* interact with a shared mechanism for initiating the defense response, the variable N-terminal domains may play a role in specific recognition. Domain swapping experiments to create chimeric proteins should address this hypothesis and define which regions of *Cf-2* and *Cf-9* are involved in ligand recognition and signal transduction.

#### Gene Duplication and the Evolution of Resistance Gene Specificities

The presence of two nearly identical functional *Cf-2* genes is in contrast to all other *R* genes so far characterized. Unlike the *Cf-9*, *N*, and *L6* genes (Jones et al., 1994; Whitham et al., 1994; Lawrence et al., 1995), transposon tagging of *Cf-2* would probably have been impossible owing to the need to inactivate both copies to bring about loss of *Cf-2* function. It seems likely that the two *Cf-2* genes arose through a recent gene duplication event, since within the open reading frames they have only diverged by three nucleotide changes. The *Cf-9*, *N*, and *L6* genes hybridize to large multigene families (although, in the case of *L6*, the multigene family maps to the homologous *M* locus), suggesting that gene duplication events are common during *R* gene evolution (Jones et al., 1994; Whitham et al., 1994; Lawrence et al., 1995; Ellis et al., 1995). Presumably, a more ancient duplication event gave rise to the MG112A and MG112B regions from a common ancestor, and since then they have diverged in *Cf0* and *Cf2*.

A diagonal plot comparing the DNA sequence of the *Cf-2* open reading frame against itself reveals a pronounced alternating pattern of A and B repeats (Figure 4D). The diagonal lines of the four consecutive B repeats (LRRs 12–15) are separated by a spacing of 72 nt (one complete repeat). However, where the A and B repeats alternate the spacing is equivalent to 144 nt (two complete repeats). The fact that these repeats are so apparent at the nucleotide level suggests that the multiple copies were generated by another recent intragenic amplification event(s). A region of such highly reiterated DNA sequences might be associated with an increase in the rate of intragenic unequal crossing over, and consequently variation could be generated at a higher rate in this domain. If such a domain encoded the ligand-binding domain of a receptor, the increased rate of evolution might be expected to generate new binding specificities. In this case, new resistance gene alleles might occasionally be created with the potential to recognize new pathogens. Further DNA sequence analysis of the *Cf-2* locus and *Cf-2* homologs in *Cf0*, *Cf2*, and *Cf5* will provide insight into the mechanisms that might generate diversity and novelty in the capacity of this *R* gene family to recognize potential ligands.

#### Experimental Procedures

##### Plant Stocks and Inoculation Procedures

The NILs of the *L. esculentum* cultivar MoneyMaker carrying either *Cf-2* (*Cf2*) or *Cf-5* (*Cf5*) and the original MoneyMaker line carrying no detectable resistance genes for *C. fulvum* (*Cf0*) were obtained from R. Oliver (University of East Anglia, Norwich, UK). Assays for disease resistance were performed exactly as described in Dickinsson et al. (1993).

##### General DNA and RNA Methods

Genomic DNA was extracted from fresh plants as described by Carroll et al. (1995), and Southern blot analysis was performed according to Dixon et al. (1995), washing in 0.5× SSC, 0.1% SDS at 65°C. RNA was prepared from young leaf material from 6-week-old glasshouse-grown tomato plants according to Scofield et al. (1992). Polyadenylated (poly[A]<sup>+</sup>) RNA was prepared by oligo(dT)-cellulose chromatography using an mRNA purification kit (Pharmacia) according to the instructions of the manufacturer.

##### Construction of Cosmid Library

Genomic DNA from plants homozygous for *Cf-2* and *Cf-9* was extracted as described by Carroll et al. (1995), except that high molecular weight DNA was recovered by spooling rather than precipitation. All subsequent techniques were carried out according to Sambrook et al. (1989), unless otherwise stated. Further purification of genomic DNA was performed on a cesium chloride-ethidium bromide gradient. This DNA was partially digested with *Mbo*I, dephosphorylated, and size fractionated twice on 10%–40% sucrose gradients. Fractionated insert DNA (1 μg) was ligated at 12°C for 18 hr in 10 μl with 250 ng of *Bam*HI-digested pCLD04541 binary cosmid vector. Ligated DNA was packaged using commercial extracts (Gigapak; Stratagene) according to the instructions of the manufacturer and transfected into SURE tet-sensitive *E. coli* (Stratagene). Over 200,000 recombinant clones were distributed into 144 pools, with each pool containing approximately 1,500 independent clones. Cosmids had an average insert size of 18 kb. Based on an estimate for the haploid tomato genome of 900 Mb, the cosmid library should represent approximately four haploid genome equivalents. To facilitate storage and screening, bacteria from each pool were harvested and used to prepare glycerol stocks and cosmid DNA. DNA from each pool was used for the preparation of slot blots for hybridization with various probes. Once the correct pool was identified, clones of interest were recovered by plating of a small proportion of the

appropriate glycerol stock, followed by conventional colony filter hybridization.

#### Tomato Transformation

All DNA for transformation into plant cells was cloned into the binary cosmid vector pCLD04541 and mobilized into *Agrobacterium tumefaciens* LBA4404. Transformation of tomato cotyledons (variety Money-maker Cf0) and plant regeneration were performed essentially as described by Fillatti et al. (1987).

#### DNA Sequencing and Analysis

All templates for sequencing were generated using a random shearing approach involving sonication and subsequent cloning into M13 (Sambrook et al., 1989). Single-stranded DNA was sequenced using the PRISM Ready Reaction Terminator Cycle Sequencing system (Applied Biosystems). Reactions were run on an Applied Biosystems 373 DNA sequencer, and data were assembled using UNIX versions of the Staden Programs Package including TED and XBAP run on a DEC 3000. Analysis of primary sequences was performed using Genetics Computer Group programs (University of Wisconsin).

#### Acknowledgments

Correspondence should be addressed to J. D. G. J. We thank R. Oliver for providing the *C. fulvum* race 4 GUS and for the Cf0, Cf2, and Cf5 NILs of tomato and P. De Wit for providing all other *C. fulvum* cultures (race 5, race 2,5, race 2,5,9, and race 5,9). We also thank C. Gebhardt, S. D. Tanksley, and M. W. Ganal for the provision of the RFLP probes. We are also grateful to O. Kolade and A. Hemmings (University of East Anglia, United Kingdom) for helpful comments regarding the possible structure of the Cf-2 LRRs and to K. E. Hammond-Kosack for constructive criticisms of this paper. We thank P. Bovill for running the ABI373 automated sequencer and S. J. Perkins, M. M. Shailer, and J. P. Darby for plant care. This work was supported by the Gatsby Foundation and Biotechnology and Biological Sciences Research Council Plant Molecular Biology Program grant PMB/523.

Received November 3, 1995; revised December 26, 1995.

#### References

Altschul, S.F., Gish, W., Millar, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Ashfield, T.A., Hammond-Kosack, K.E., Harrison, K., and Jones, J.D.G. (1994). Cf gene-dependent induction of a  $\beta$ -1,3-glucanase promoter in tomato plants infected with *Cladosporium fulvum*. *Mol. Plant Microbe Interact.* **7**, 645–657.

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.

Bowling, S.A., Guo, A., Cao, H., Gordon, S., Klessig, D.F., and Dong, X. (1994). A mutation in *Arabidopsis* that leads to constitutive expression of systemic acquired resistance. *Plant Cell* **6**, 1845–1857.

Carroll, B.J., Klimyuk, V.I., Thomas, C.M., Bishop, G.J., Harrison, K., and Jones, J.D.G. (1995). Germinal transpositions of the maize element *Dissociation* from T-DNA loci in tomato. *Genetics* **139**, 407–420.

Chang, C., Schaller, E., Patterson, S.E., Kwok, S.F., Meyerowitz, E.M., and Bleecker, A.B. (1992). The TMK1 gene from *Arabidopsis* codes for a protein with structural and biochemical characteristics of a receptor protein kinase. *Plant Cell* **4**, 1263–1271.

Dickinson, M., Jones, D.A., and Jones, J.D.G. (1993). Close linkage between the Cf-2/Cf-5 and *Mf* resistance loci in tomato. *Mol. Plant Microbe Interact.* **6**, 341–347.

Dietrich, R.A., Delaney, T.P., Uknes, S.J., Ward, E.R., Ryals, J.A., and Dangl, J.L. (1994). Arabidopsis mutants simulating disease resistance response. *Cell* **77**, 565–577.

Dixon, M.S., Jones, D.A., Hatzixanthis, K., Ganal, M.W., Tanksley,

S.D., and Jones, J.D.G. (1995). High resolution mapping of the physical location of the tomato Cf-2 gene. *Mol. Plant Microbe Interact.* **8**, 200–206.

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.

Fillatti, J.J., Kiser, J., Rose, R., and Comai, L. (1987). Efficient transfer of a glyphosate tolerance gene into tomato using a binary *Agrobacterium tumefaciens* vector. *Biotechnology* **5**, 726–730.

Flor, H.H. (1946). Genetics of pathogenicity in *Melampsora lini*. *J. Agr. Res.* **73**, 335–357.

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.

Greenberg, J.T., Guo, A., Klessig, D.F., and Ausubel, F.M. (1994). Programmed cell death in plants: a pathogen-triggered response activated coordinately with multiple defense functions. *Cell* **77**, 551–563.

Hammond-Kosack, K.E., and Jones, J.D.G. (1994). Incomplete dominance of tomato Cf genes for resistance to *Cladosporium fulvum*. *Mol. Plant Microbe Interact.* **7**, 58–70.

Hammond-Kosack, K.E., Jones, D.A., and Jones, J.D.G. (1994). Identification of two genes required in tomato for full Cf-9-dependent resistance to *Cladosporium fulvum*. *Plant Cell* **6**, 361–374.

Johal, G.S., Hulbert, S.H., and Briggs, S.P. (1995). Disease lesion mimics of maize: a model for cell death in plants. *Bioessays* **17**, 685–692.

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.

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.

Jones, J.D.G. (1994). Paranoid plants have their genes examined. *Curr. Biol.* **4**, 749–751.

Joosten, M.H.A.J., Cozijnsen, T.J., and De Wit, P.J.G.M. (1994). Host resistance to a fungal tomato pathogen lost by a single base-pair change in an avirulence gene. *Nature* **367**, 384–386.

Kobe, B., and Deisenhofer, J. (1993). Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature* **366**, 751–756.

Kobe, B., and Deisenhofer, J. (1994). The leucine-rich repeat: a versatile binding motif. *Trends Biochem. Sci.* **19**, 415–421.

Lamb, C.J. (1994). Plant disease resistance genes in signal perception and transduction. *Cell* **76**, 419–422.

Langford, A.N. (1948). Autogenous necrosis in tomatoes immune from *Cladosporium Cooke*. *Can. J. Res. (C)* **26**, 35–64.

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.

L Levine, A., Tenhaken, R., Dixon, R., and Lamb, C. (1994). H<sub>2</sub>O<sub>2</sub> from the oxidative burst orchestrates the plant hypersensitive disease resistance response. *Cell* **79**, 583–593.

Martin, G.B., Brommonschenkel, S.H., Chunwongse, J., Frary, A., Ganal, 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.

May, M.J., Hammond-Kosack, K.E., and Jones, J.D.G. (1996). Involvement of reactive oxygen species, glutathione metabolism and lipid peroxidation in the Cf-gene-dependent defence response of tomato cotyledons induced by race-specific elicitors of *Cladosporium fulvum*. *Plant Physiol.*, in press.

Mindrinos, 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.

Nasrallah, J.B., and Nasrallah, M.E. (1993). Pollen-stigma signalling in the sporophytic self-incompatibility response. *Plant Cell* **5**, 1325–1335.

Oliver, R.P., Farman, M.L., Jones, J.D.G., and Hammond-Kosack, K.E. (1993). Use of fungal transformants expressing  $\beta$ -glucuronidase activity to detect infection and measure hyphal biomass in infected plant tissues. *Mol. Plant Microbe Interact.* **6**, 521–525.

Salmeron, J.M., Barker, S.J., Carland, F.M., Mehta, A.Y., and Staskawicz, B.J. (1994). Tomato mutants altered in bacterial disease resistance provide evidence for a new locus controlling pathogen recognition. *Plant Cell* **6**, 511–520.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Scofield, S., Harrison, K.A., Nurrish, S.J., and Jones, J.D.G. (1992). Promoter fusions to the *Ac* transposase gene confer distinct patterns of *Ds* somatic and germinal excision in tobacco. *Plant Cell* **4**, 573–582.

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

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.

Stevens, M.A., and Rick, C.M. (1988). Genetics and breeding. In *The Tomato Crop*, J.G. Atherton and J. Rudich, eds. (London: Chapman and Hall), pp. 35–109.

van Kan, J.A.L., van Den Ackerveken, G.F.J.M., and De Wit, P.J.G.M. (1991). Cloning and characterization of cDNA of avirulence gene *avr9* of the fungal pathogen *Cladosporium fulvum*, causal agent of tomato leaf mold. *Mol. Plant Microbe Interact.* **4**, 52–59.

Vera-Estrella, R., Blumwald, E., and Higgins, V.J. (1992). Effect of specific elicitors of *Cladosporium fulvum* on tomato suspension cells: evidence of the involvement of active oxygen species. *Plant Physiol.* **99**, 1208–1215.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* **14**, 4683–4690.

Walker, J.C. (1994). Structure and function of receptor-like protein kinases of higher plants. *Plant Mol. Biol.* **26**, 1599–1609.

Ward, E.R., Uknes, S.J., Williams, S.C., Dincher, S.S., Wiederhold, D.L., Alexander, D.C., Ahl-Goy, P., Métraux, J.-P., and Ryals, J.A. (1991). Coordinate gene activity in response to agents that induce systemic acquired resistance. *Plant Cell* **3**, 1085–1094.

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.

#### GenBank Accession Numbers

The accession numbers for the sequences reported in this paper are U42444 for *Cf-2.1* and U42445 for *Cf-2.2*.