

## REVIEW

# Clusters of Resistance Genes in Plants Evolve by Divergent Selection and a Birth-and-Death Process

**Richard W. Michelmore<sup>1</sup> and Blake C. Meyers**

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

Classical genetic and molecular data show that genes determining disease resistance in plants are frequently clustered in the genome. Genes for resistance (*R* genes) to diverse pathogens cloned from several species encode proteins that have motifs in common. These motifs indicate that *R* genes are part of signal-transduction systems. Most of these *R* genes encode a leucine-rich repeat (LRR) region. Sequences encoding putative solvent-exposed residues in this region are hypervariable and have elevated ratios of nonsynonymous to synonymous substitutions; this suggests that they have evolved to detect variation in pathogen-derived ligands. Generation of new resistance specificities previously had been thought to involve frequent unequal crossing-over and gene conversions. However, comparisons between resistance haplotypes reveal that orthologs are more similar than paralogs implying a low rate of sequence homogenization from unequal crossing-over and gene conversion. We propose a new model adapted and expanded from one proposed for the evolution of vertebrate major histocompatibility complex and immunoglobulin gene families. Our model emphasizes divergent selection acting on arrays of solvent-exposed residues in the LRR resulting in evolution of individual *R* genes within a haplotype. Intergenic unequal crossing-over and gene conversions are important but are not the primary mechanisms generating variation.

Plants, like animals, are continually challenged by a myriad of potential pathogens. There is increasing evidence that defense systems of plants are at least as complex as vertebrate defense systems. Unlike animals, however, plants do not have a circulatory system and therefore cannot rely on a specialized, proliferative immune system. Each plant cell has to be capable of defense, even though this defense is coordinated locally and systemically between cells. There are a variety of types of resistance genes and mechanisms, some induced and some constitutive (for review, see Godiard et al. 1994; Michelmore 1995; Hammond-Kosack and Jones 1997). Often, although not always, disease resistance in plants is determined by single, usually dominant, genes. The recent cloning of several such resistance genes (*R* genes) is providing insight into their function and evolution. The defense system of plants may be ancient and predate the evolution of the immune system; genes similar to plant *R* genes have been identified in mammals (Hammond-Kosack and Jones 1997; van der Biezen and Jones 1998).

In this review we consider what is known about

the genomic organization and evolution of disease resistance genes in plants. The picture that is emerging for the organization and evolution of plant *R* genes is similar to that of the vertebrate major histocompatibility complex (MHC), T-cell receptor (TCR), and immunoglobulin genes. Therefore, although the specific types of genes involved are different, the evolutionary forces shaping the plant and vertebrate defense systems may be similar. We propose a model for the evolution of plant *R* genes that is adapted and expanded from a model developed for the evolution of vertebrate MHC and immunoglobulin gene families. Definitions of terminology can be found at the end of the text.

## Classical Genetics

Resistance to many diseases, particularly those caused by biotrophic fungal diseases, is determined by individual members of families of dominant genes, each member conferring resistance to a specific strain of the pathogen. When parallel genetic studies have been made, specificity has often been shown to be conferred by a "gene-for-gene" interaction; for every gene for resistance in the host there is a matching gene for avirulence in the pathogen

<sup>1</sup>Corresponding author.  
E-MAIL [rwmichelmore@ucdavis.edu](mailto:rwmichelmore@ucdavis.edu); FAX (530) 752-9659.

## MICHELMORE AND MEYERS

(Flor 1956). The gene-for-gene interaction has now been demonstrated or inferred in >20 diseases caused by taxonomically diverse pathogens and probably functions in many more (Flor 1971; Crute 1986). The gene-for-gene interaction is undoubtedly an oversimplification; however, it is a useful predictive genetic description of the interaction between plants and their pathogens.

Classical genetics has increasingly demonstrated that resistance genes tend to be clustered in the genome. Early studies were often limited to the study of a single pathogen. However, the advent of comprehensive genetic maps based on molecular markers has allowed genes for resistance to multiple diseases to be readily mapped relative to each other. *R* loci may be single genes with multiple alleles (Table 1). The *L* locus in flax (13 alleles) and *RPM1* in *Arabidopsis* (two alleles, presence or absence) are of this type. More commonly, *R* genes are organized in clusters that show varying levels of recombination between the component genes (Table 1). Genes within a single cluster can determine resistance to very different pathogens. On the basis of their clustered distribution and by inference from other cell-cell recognition systems, *R* genes were hypothesized to encode functionally and evolutionarily related members of recognition systems.

## Molecular Evidence

Molecular data have supported this hypothesis. Nearly 20 *R* genes have now been cloned. Their isolation, structure, and potential function have been reviewed extensively (e.g., Staskawicz et al. 1995; Bent 1996; Baker et al. 1997; Hammond-Kosack and Jones 1997) and only features relevant to the evolution of *R* genes will be discussed here. All except two (*Hm1* and *mlo*) seem to encode components of signal transduction systems. *R* genes encode a number of protein motifs in a variety of combinations (Fig. 1; Table 2) that are characteristic of receptors in yeast, *Drosophila* and vertebrates. The most prevalent class contains a nucleotide-binding site (NBS) and a leucine-rich repeat (LRR) region. Some are composed predominantly of an LRR region and transmembrane (TM) domain. Another is comprised of a LRR region and a protein kinase domain. Finally, one (*Pto*) is comprised of only a protein kinase domain but requires a NBS-LRR gene (*Prf*) for function. *R* genes therefore seem to encode receptors that detect the presence of the pathogen directly or indirectly. The signal passes through a network of signal-transduction cascades and a somewhat generic resistance response is elicited that may

or may not involve host cell death (for review, see Dangl et al. 1996; Hammond-Kosack and Jones 1997).

Restriction fragment length polymorphism (RFLP) analysis with most cloned resistance genes reveal clusters of related sequences, even when only a single resistance specificity had been detected by classical genetics (Table 2). *L* of flax is an interesting exception; the 13 specificities known at the *L* locus are truly allelic. *M*, the presumed homologous locus in this ancient tetraploid, has ~86% nucleotide identity to *L* and comprises an array of ~15 genes (paralogs); seven specificities are known at *M* (Ellis et al. 1997). Occasionally two functional sequences encode the same specificity, as in the case of *Cf2* (Dixon et al. 1996). More usually only a single gene in the cluster determines resistance as shown by mutant or transgenic complementation analysis. The genetic and hybridization data therefore indicate that resistance genotypes should be considered as haplotypes rather than individual genes.

Molecular data is also revealing that clusters of *R* genes may contain sequences related in function but not sequence. *Prf*, an NBS-LRR gene, is within a cluster of five *Pto* homologs that encode protein kinases; both genes are required for resistance to *Pseudomonas syringae* pv. *tomato* (Salmeron et al. 1996). *Mi*, an NBS-LRR gene, is linked loosely with *Cf2*, a LRR-TM gene. Therefore, genetic linkage of phenotypically defined *R* genes does not necessarily imply sequence similarity.

There are a large number of sequences with similarity to *R* genes in plant genomes. PCR using degenerate primers designed to amplify sequences conserved between *R* genes has allowed identification of families of sequences from several plants (Kanazin et al. 1996; Leister et al. 1996; Yu et al. 1996; Lagudah et al. 1997; Leister et al. 1998). It is now facile to clone homologs but it remains difficult to prove individual specificities (Michelson 1995). Large-scale sequencing of random expressed sequence tags (ESTs) and genomic clones has also identified numerous *R* gene homologs (Botella et al. 1997; Bevan et al. 1998). Extrapolating from current data, it seems likely that *Arabidopsis* has >400 resistance gene candidates (~2% of its genes) and plants with larger genomes may have significantly greater numbers. Among the unanswered questions are: How many of these *R* gene candidates encode functional resistance genes? Is there a cost to expressing many *R* genes? What is the molecular basis of specificity? How do plants evolve new recognition specificities to keep pace with rapidly changing pathogen populations? What are the evolutionary forces shaping the size and structure of *R* gene clusters?

## CLUSTERS OF RESISTANCE GENES

**Table 1. Clusters of Plant *R* Genes Identified by Classical Genetics**

Species	<i>R</i> -gene locus	Linked resistance specificities	No. of pathogens detected	Size (cM)	References
<i>Arabidopsis</i>	MRC-F	6 <i>RPP</i> genes, <i>RAC2</i>	2	10+	Kunkel (1996); Holub (1997)
<i>Arabidopsis</i>	MRC-H	5 <i>RPP</i> genes, <i>RPS2</i> , <i>TOM1</i>	3	15 ±	Kunkel (1996); Holub (1997)
<i>Arabidopsis</i>	MRC-J	5 <i>RPP</i> genes, <i>RAC3</i> , <i>RPS4</i> , <i>HRT1</i> , <i>TTR1</i>	5	15 ±	Kunkel (1996); Holub (1997)
Barley	<i>Mla</i>	18 <i>Mla</i> genes	1	0	Jahoor et al. (1993); Jørgenson (1994)
Flax	<i>L</i>	13 <i>L</i> alleles	1	0	Islam and Shepherd (1991)
Flax	<i>M</i>	7 <i>M</i> genes	1	<1	Flor (1965); Mayo and Shepherd (1980)
Flax	<i>N</i>	<i>N</i> , <i>N1</i> , <i>N2</i>	1	<1	Islam and Shepherd (1991); Ellis et al. (1997)
Flax	<i>P</i>	7 <i>P</i> genes	1	0	Islam and Shepherd (1991); Ellis et al. (1997)
Lettuce	<i>Dm1/3</i>	8+ <i>Dm</i> genes, <i>Ra</i>	2	20	Farrara et al. (1987); Kesseli et al. (1994)
Lettuce	<i>Dm5/8</i>	4 <i>Dm</i> genes, <i>Tu</i> , <i>Mo2</i> , <i>plr</i>	4–5	9	Farrara et al. (1987); Witsenboer et al. (1995); Kesseli et al. (1994)
Lettuce	<i>Dm4/7</i>	5+ <i>Dm</i> genes	1	0	Farrara et al. (1987)
Maize	<i>Rp1</i>	14 <i>Rp1</i> genes (A–N), <i>Rp5</i> , <i>Rp6</i> , <i>Rpp9</i>	2	3	Saxena and Hooker (1968); Hooker (1985); Bennetzen et al. (1991)
Maize	<i>Rp3</i>	6 <i>Rp3</i> genes (A–F)	1	N.R.	Saxena and Hooker (1974); Sanz-Alferez et al. (1995)
Potato	chromosome 5	<i>R1</i> , <i>Pi01–QTL</i> , <i>Rx2</i> , <i>Gpa</i>	3	N.R.	Leister et al. (1996)
Rice	chromosome 11	4+ <i>Xa</i> genes, 5+ <i>Pi</i> genes, 1 QTL	2	30	Yoshimura et al. (1983); Mackill and Bonman (1992); Kinoshita (1993)
Soybean	linkage group F	<i>Rsv1</i> , <i>Rpv</i> , <i>Rps3</i>	3	N.R.	Diers et al. (1992); Yu et al. (1994)
Soybean	linkage group J	<i>Rps2</i> , <i>Rmd</i> , <i>Rj2</i> , 1 QTL	3–4	3.8	Polzin et al. (1994); Kanazin et al. (1996)
Tomato	<i>Cf2</i> , <i>Cf5</i>	<i>Cf2</i> , <i>Cf5</i> , <i>OI-1</i> , <i>TY-1</i> , <i>Mi</i> , <i>Meu-1</i>	5	N.R.	Dickinson et al. (1993); Jones et al. (1993); Van der Beek et al. (1994); Zamir et al. (1994); Kaloshian et al. (1995)
Tomato	<i>Cf4/9</i>	<i>Cf1</i> , <i>Cf4</i> , <i>Cf9</i>	1	0	Kerr and Bailey (1964); Jones et al. (1993)

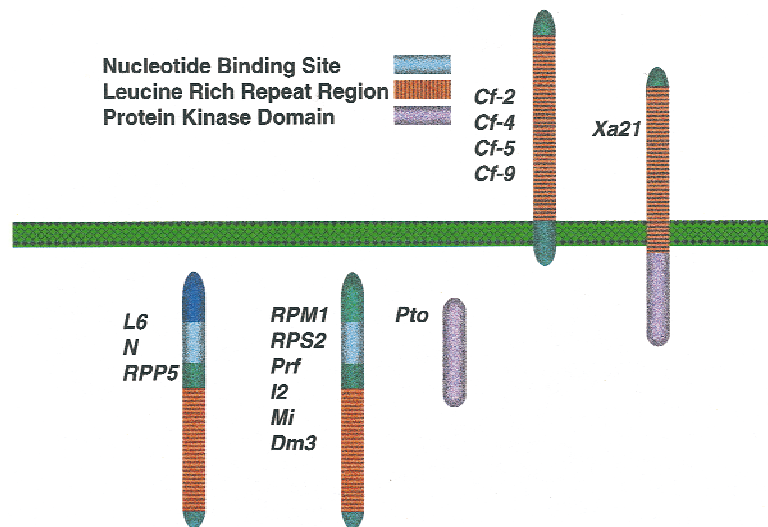
(N.R.) Not reported; usually because genes segregated in different crosses.

**Basis of Specificity**

The basis of specificity is unknown for most *R* genes. *R*-gene products are postulated to have receptor and effector domains (Fig. 2; Hammond-Kosack and Jones 1997) and patterns of nonsynonymous versus synonymous nucleotide-substitution ratios support this hypothesis (see below). A variety of motifs sug-

gest involvement in protein–protein interactions; these could be involved in ligand binding or effector functions. Interactions between the putative components of a receptor complex have not been demonstrated yet. Mutations to losses of resistance are not particularly informative as to which domains determine specificity because mutations anywhere within the receptor or effector domains could abol-

## MICHELMORE AND MEYERS



**Figure 1** Major protein motifs shared between the deduced products of cloned resistance genes.

ish function. Various mutant alleles of *Prf*, *RPS2*, *RPM1*, and *Pto* encode nonfunctional proteins that differ from wild type by a variety of single amino acids (Bent et al. 1994; Mindrinos et al. 1994; Grant et al. 1995; Salmeron et al. 1996; Scofield et al. 1996). Domain swaps between closely related proteins are likely to be much more instructive.

*Pto* is the best-characterized *R* gene. It is a small gene and diverse protein kinases have been extensively characterized structurally and biochemically. The product of the bacterial avirulence gene, *AvrPto*, interacts directly with the *Pto* gene product in the yeast two-hybrid system. The phenotypes of mutant and recombinant *Pto* genes in the yeast two-hybrid analysis correlate with biological activity in planta (Scofield et al. 1996; Tang et al. 1996). Domain swaps between *Pto* and the closely related paralog, *Fen*, identified a few amino acids in the ligand-binding domain (protein-kinase domain seven) as being critical for specificity (Scofield et al. 1996; J. Rathjen, J. Chang, D. Lavelle, B. Staskawicz, and R.W. Michelson, unpubl.).

The LRR motif is a common motif that is thought to be involved in protein-protein interactions and known to be involved in ligand binding in porcine ribonuclease inhibitor (PRI; Kobe and Deisenhofer 1994, 1995). In PRI, the residues corresponding to the hypervariable residues in *R* gene products (see below) are part of a  $\beta$ -strand/ $\beta$ -turn structure of the LRR with a consensus sequence xx-LxLxx. The conserved leucines (L) project into the hydrophobic core, whereas the other residues (x) form a solvent-exposed surface that is involved in ligand binding (Kobe and Deisenhofer 1995). The

solvent-exposed surfaces of each LRR are arranged as a curved array; PRI accommodates its ligand by flexing along its backbone. In *R* genes, the conserved positions in the consensus sequence contain a variety of aliphatic residues (a; Fig. 2). *R* genes are unlikely to have as regular a structure as PRI because the amino acids in the backbone are more variable and there is less evidence that they form regular  $\alpha$ -helices (Hammond-Kossack and Jones 1997; Jones and Jones 1997). Three-dimensional modeling suggests that the thyrotropin- and choriogonadotropin-receptor domains and LRRs of other proteins have less regular LRR structures but still comprise arrays of  $\beta$ -strand/ $\beta$ -turn structures (Kajava et al. 1995; Jiang et al. 1995; Kajava 1998).

Genetic rather than biochemical evidence currently provides the strongest evidence that the LRR interacts with a pathogen-derived ligand. Binding of an avirulence gene product or a plant protein to an LRR region has yet to be reported for any *R* gene product. Allelic comparisons and domain swaps between alleles of *L* indicated that specificity was determined by the 3' LRR-encoding part of the gene (Ellis et al. 1997). Sequences of the 5' end of the LRR-encoding region of *Cf* genes are more variable than the 3' regions (Parniske et al. 1997; Thomas et al. 1997). Detailed analysis of 11 *Cf* homologs revealed that in the first 16 LRRs, the amino acids in the xx(a)x(a)xx part of the LRR consensus were hypervariable, and the ratios of non-synonymous to synonymous substitutions ( $K_A:K_S$  ratio) of the nucleotides encoding these residues (excluding the conserved aliphatic residues) were greater than one; therefore these residues may be under diversifying selection (Parniske et al. 1997). All types of LRR-encoding *R* genes exhibit a similar pattern (Table 3; Meyers et al. 1998b). The NBS region appears to be under purifying selection consistent with its proposed but unproven effector function. The LRR region, particularly the carboxy-terminal half of the LRR region encoded by *Dm3* homologs in lettuce, contains an alternating pattern of hypervariable amino acids in the xx(a)x(a)xx consensus and intervening stretches of conserved residues that are predicted to be structural. The  $K_A:K_S$  ratios for the nucleotides encoding these hypervariable residues are significantly greater than one, suggesting that there is an advantage to high amino acid diversity in this region. Therefore, both the amino-terminal end of the probably extracyto-

## CLUSTERS OF RESISTANCE GENES

**Table 2. Clusters of Plant *R* Genes Characterized by Molecular Analysis**

<i>R</i> -gene locus	Structural subclass <sup>a</sup>	Species	Members of gene family <sup>b</sup>	Size of locus	References
<i>Cf2/Cf5</i>	LRR-TM	tomato	2+	N.R.	Jones et al. (1993); Dixon et al. (1996); Hammond-Kosack and Jones (1997)
<i>Cf4/Cf9</i>	LRR-TM	tomato	5	36 kb	Jones et al. (1993); Parniske et al. (1997); Thomas et al. (1997)
<i>Dm3</i>	NBS-LRR	lettuce	24+	4 ± Mb	Farrara et al. (1987); Meyers et al. (1998a)
<i>Dm13</i>	NBS-LRR	lettuce	13+	N.R.	Meyers (1998)
<i>I2</i>	NBS-LRR	tomato	4+	N.R.	Ori et al. (1997)
<i>M</i>	TIH-NBS-LRR	flax	15±	<1 Mb	Anderson et al. (1997); Ellis et al. (1997)
<i>N</i>	TIH-NBS-LRR	tobacco	4+	N.R.	Whitham et al. (1994)
<i>Pto</i>	kinase	tomato	5, plus <i>Prf</i>	65 ± kb	D. Lavelle and R. Michelmore, (unpubl.)
<i>RPP5</i>	TIH-NBS-LRR	<i>Arabidopsis</i>	7	50 ± kb	Parker et al. (1996); Dangl and Holub (1997)
<i>Xa21</i>	LRR-kinase	rice	8+	230 ± kb	Ronald et al. (1992); Song et al. (1995)

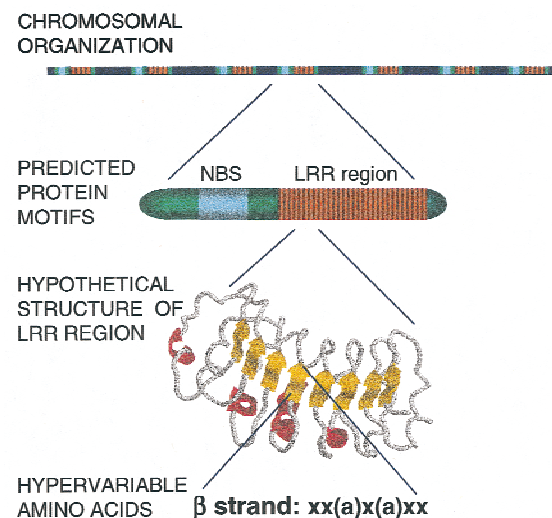
(N.R.) not reported.

<sup>a</sup>(TIH) Toll-interleukin homology region; (NBS) nucleotide binding site; (LRR) leucine-rich repeats; (TM) transmembrane domain.<sup>b</sup>Members of the gene family at each locus were determined in a single genotype by mapping of RFLPs resulting from the hybridization of the cloned resistance gene from that locus or by genomic sequencing of the region containing the cluster. This number includes pseudogenes.

plasmic LRR regions of the LRR-TM type *R* gene products and the carboxy-terminal end of the predicted cytoplasmic LRR regions of the NBS-LRR type *R* gene products show an alternating pattern of hypervariable and conserved residues. The LRR region encoded by *Xa21* and related paralogs show the same pattern. Regardless of the precise structure, the alternating pattern of variation and high  $K_A:K_S$  ratios are evidence that these LRRs have a series of contact points for ligand binding (Fig. 2) that seem to be under diversifying selection.

A structure for LRR regions with arrays of potential ligand-binding surfaces has several implications for *R*-gene function. The most important implications are the extremely large number of binding specificities that could be encoded by groups of genes with such arrays (Fig. 2) and the ease with which new binding specificities could be generated by recombination and gene conversion. In addition to different combinations of LRRs providing different binding characteristics, variation in amino acids in the backbone between the hypervariable regions might change the relative orientations of the  $\beta$ -strands providing another level of variation for binding specificity. The number of repeats does not seem to be highly conserved; this can vary from 14 to >40 (Jones and Jones 1997); *Dm3* homologs encode at least 41 repeats (Meyers et al. 1998b). *Cf2* differs from *Cf5* by precisely six LRRs (Hammond-Kosack and Jones 1997). Different functional alleles of *L* vary in the length of their encoded LRR region

(Ellis et al. 1997). Furthermore, LRR regions are large relative to predicted avirulence gene products. It is therefore unlikely that all LRRs within a particular protein are involved in binding each avirulence gene product. Different binding specificities could



**Figure 2** The organization and structure of resistance genes and their products. The color coding of the protein motifs is the same as for Figure 1. The hypothetical structure shown for a portion of the LRR region is prepared from a model for the thyrotropin receptor (Kajava et al. 1995). In the consensus sequence for the  $\beta$ -strand, (x) any amino acid; (a) aliphatic residues.



## MICHELMORE AND MEYERS

**Table 3. Ratio of Nonsynonymous to Synonymous Nucleotide Substitutions ( $K_A:K_S$ ) in Different Regions of Plant *R* Genes**

Species	Genes	NBS region	LRR: intervening	LRR: xx(a)x(a)xx
Lettuce	<i>RG2</i> family	0.37	0.63 <sup>a</sup>	2.06 <sup>a</sup>
Tomato	<i>I2C-1</i> vs <i>I2C-2</i>	0.47	0.58	1.17
Tomato	<i>Mi</i> copy 1 vs copy 2	0.43	0.76	1.93
Flax	<i>L6</i> vs <i>M</i>	0.44	0.83	1.57
Tomato	<i>Cf4/Cf9</i> homologs		0.59 <sup>b</sup>	1.32 <sup>b</sup>
Rice	<i>Xa21</i> copies <i>B</i> , <i>D</i> , <i>F</i>		0.52	2.11

<sup>a</sup>For the 3' LRR region.<sup>b</sup>For the 5' LRR region (Parniske et al. 1997).

be generated by interactions with different LRRs within a protein, and a single protein could interact with a variety of ligands. *RPM1* determines resistance against bacteria expressing two very different *Avr* genes (Grant et al. 1995). The tomato *Mi* gene confers resistance to both nematodes and aphids (Rossi et al. 1998). However, it has yet to be demonstrated in these cases whether the same or different molecules are detected by the plant *R* gene.

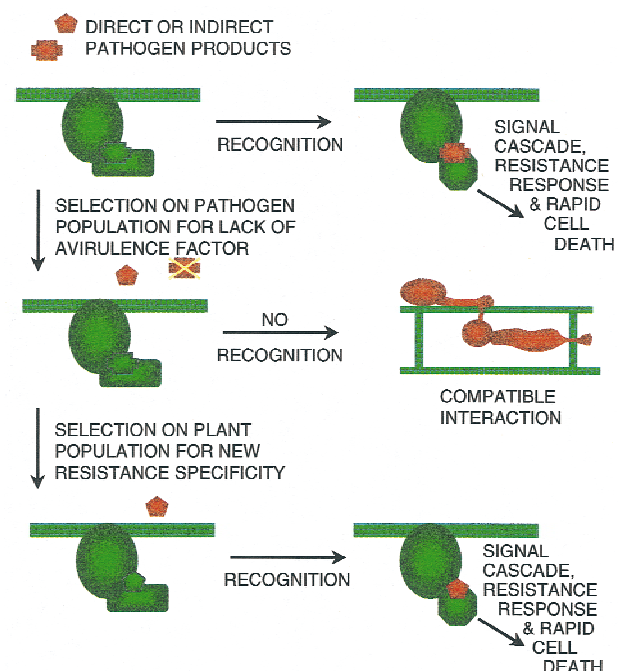
Post-transcriptional and post-translational modifications add further levels of potential complexity that have yet to be studied for *R* gene products. The prevalence and significance of the alternative mRNA splicing that seems to occur with *N*, *L6*, and *RPP5* is unknown (Whitham et al. 1994; Lawrence et al. 1995; Parker et al. 1997). Amino acid differences between *Cf* gene products alter potential glycosylation sites (Jones and Jones 1997). For several mammalian extracellular receptors, the pattern of N-glycosylation influences ligand binding (Zhang et al. 1995).

#### Mechanisms Generating Variation in Specificity

Cycles of detection and mutation characterize the coevolution of plants and potential pathogens (Fig. 3). This is particularly evident with the specialized airborne pathogens, such as the mildews and rusts, giving rise to a boom-and-bust cycle (Suneson 1960) for the efficacy of resistance genes. However, the genetic events underlying the changes in the plant and pathogen are understood poorly.

Instability of some resistance loci, especially *Rp1* of maize, led to the idea that clusters of resistance genes were inherently unstable, fast-evolving complexes. This instability is associated with recombination, and/or deletions (Sudupak et al. 1993; Anderson et al. 1996). Susceptible recombinants at

*Rp1* had both possible combinations of flanking alleles (Richter et al. 1995). Unequal crossing-over following meiotic mispairing and to a lesser extent gene conversion have been invoked as the major ways in which novel resistance specificities are generated (Pryor and Ellis 1993; Richter et al. 1995; Hammond-Kosack and Jones 1997; Hulbert 1997; Parniske et al. 1997). Such processes are almost certainly the primary mechanisms for concerted evolution of large multigene families such as vertebrate globins in which large amounts of homogeneous product are required (Scott et al. 1984). However, the roles of unequal crossing-over and gene conver-

**Figure 3** Alternating cycle of selection during coevolution of plant and pathogen.

## CLUSTERS OF RESISTANCE GENES

sion in the evolution of other multigene families are less clear and more controversial (Ohta 1991; Li 1997; Nei et al. 1997).

The assumption has been that resistance-gene complexes are dynamic, unstable, fast-evolving haplotypes and that high rates of unequal crossing-over and gene conversion are required to keep pace with changes in pathogen populations. However, high rates of unequal crossing-over and gene conversion would produce rapid divergence between haplotypes of different species. There would be high levels of polymorphism for both the number and sequence of *R* genes between haplotypes. Concerted evolution would tend to homogenize genes within a haplotype. There would be no obvious allelic relationship between genes in different haplotypes. Paralogs would be more similar than orthologs. In contrast to the above expectations, the recent comparisons of haplotype structure have been particularly informative and led to a very different picture. Studies on the *Pto*, *Dm*, and *Cf* clusters reveal an initially surprisingly stable picture with orthologs being more similar than paralogs.

The structures of three haplotypes of the *Pto* cluster in tomato have now been determined (D. Lavelle and R. Michelmore, unpubl.). One was the resistant haplotype that had been introgressed from *Lycopersicon pimpinellifolium* and two were susceptible haplotypes from *Lycopersicon esculentum*. Four to six *Pto* homologs are present within 65 kb along with a single copy of *Prf*. The closest relative of each gene is an ortholog rather than a paralog resulting in an obvious allelic relationship between homologs in different haplotypes. Only one pseudogene was apparent in the sequence of one haplotype as well as two partial gene segments. The cluster has evolved by a series of ancient duplications and more recent deletions. Unequal crossing-over with exchanges both between and within coding regions has occurred but has been infrequent.

The major cluster of resistance genes in lettuce is comprised of over 24 *R* gene candidates (RGCs) spanning several megabases (Meyers et al. 1998a). Genetic and mutation data indicate that one RGC2 sequence encodes *Dm3*. The sequences within a genotype vary from 55%–95% amino acid identity. This is considerably greater than the differences observed between RGCs from different haplotypes and species (S.-S. Woo and R. Michelmore, unpubl.). The majority of RGCs sequenced appear to be functional and expressed, although some are clearly pseudogenes. The numbers of RGCs vary between genotypes. Several spontaneous losses of resistance have resulted from deletions but the involvement of

RGCs at the breakpoints has not been demonstrated (D. Chin and R. Michelmore, unpubl.).

Three *Cf4/9* haplotypes that originated from different tomato species have been sequenced (Parniske et al. 1997). The haplotype with no known resistance specificities contained only a single *R*-gene homolog. The other two haplotypes spanned ~35 kb and contained five paralogs, all oriented in the same direction. In all three haplotypes, the *R*-gene cluster was flanked by convergently oriented genes with similarity to plant lipoxygenases (*Lox*). In addition, the majority of the *R*-gene homologs within each cluster were flanked by partial *Lox* sequences indicating that *Lox* sequences had been duplicated along with the *R* gene. There were no other genes located within the cluster. Extensive blocks of sequence similarity were detected between the paralogs. The lengths and positions of these blocks varied greatly; however, their order was not changed indicating evolution by deletion following duplication of the *R* gene. There were no obvious *R* pseudogenes, which was in contrast to the large amount of intergenic rearrangements observed. Some homologs from different haplotypes had almost identical sequences indicating little sequence exchange between paralogs. Other homologs seemed to be a patchwork of short stretches of sequence similarity indicating the involvement of recombination or gene conversion.

Only a single haplotype of the *Xa21* cluster in rice has been partially characterized at the sequence level (Song et al. 1997). This revealed two subfamilies of paralogs. However, the subfamilies had a mixture of sequences 5' to the open reading frame (ORF) suggesting that there was a hot-spot for recombination just 3' to the start codon and exchange of promoter regions between paralogs.

The sequencing of more clusters from multiple genotypes will reveal the relative orientations of genes, the numbers of pseudogenes, the degree of sequence divergence in intergenic regions, and the relationship between physical position and sequence similarity of the genes. This will provide additional data on the frequency of gene conversion tracts and insertions/deletions characteristic of unequal crossing-over events.

#### Birth-and-Death Model for Evolution of Resistance-Gene Clusters

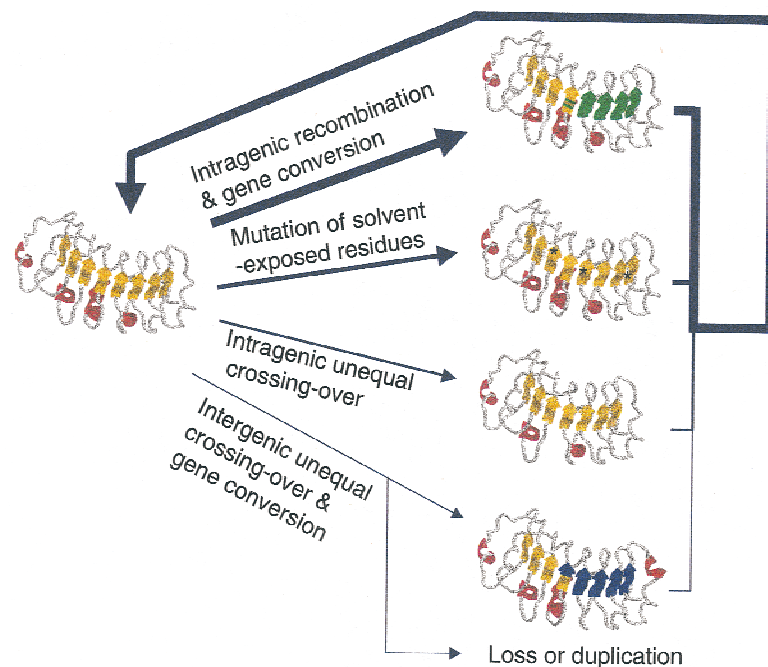
A model is required that takes into account the probable structure of *R*-gene products and accommodates the observed stability of sequences between haplotypes as well as the genetic data on in-

## MICHELMORE AND MEYERS

stability, yet allowing the rapid generation of new recognition specificities. Although unequal crossing-over is clearly occurring at a low frequency, it is not occurring at a high enough rate to homogenize sequences and it is probably too infrequent to generate new specificities to allow coevolution with pathogen populations. Frequent unequal crossing-over and gene conversion might actually be counterproductive to generating and maintaining variation. Therefore, we propose that NBS-LRR-type *R* genes are evolving mainly by divergent evolution of individual genes and a birth-and-death process (Fig. 4) similar to that envisaged for MHC and immunoglobulin genes (Nei et al. 1997).

Salient features of the model include the following:

1. The majority of changes in specificity are caused by interallelic recombination and gene conversion that alter the combinations and/or orientations of the arrays of solvent-exposed residues in the LRR region. Recombination and gene conversion between paralogs is rare.
2. Further changes in specificity result either from mutations in the solvent-exposed region of the LRRs caused by random nucleotide changes or from interallelic unequal crossing-over or gene conversion because of mispairing within the LRR region.
3. Continued mutation and interallelic recombination will result in the selection of variants encoding increasingly effective resistance genes (increased ligand binding affinities?).
4. Sequences with advantageous ligand-binding characteristics will increase in the population. Multiple specificities are maintained in the population because of frequency-dependent selection and hitch-hiking caused by selection acting on the whole cluster of resistance genes.
5. Rare unequal crossing-over events result in duplications and deletions of either single genes or blocks of genes as well as occasional formation of chimeras between paralogs.
6. Recently duplicated sequences are unstable because of a high degree of sequence similarity resulting in relatively frequent unequal crossing-over. This leads to further duplications and deletions.
7. Rapid divergence of intergenic regions (and possibly introns) reduces the frequency of unequal crossing-over. Variants and derivatives become fixed in the haplotype.
8. Duplicated genes diverge (1 and 2 above). Some will have altered ligand-binding characteristics; others will become pseudogenes.



**Figure 4** Consequences of genetic changes in resistance genes on the LRR region. Only a portion of the hypothetical structure of the LRR region is shown (see Fig. 2).

## Evidence for Model and Implications

Clearly some components of the model remain unproven. However, the model is consistent with current data and does reconcile the maintenance of sequence diversity within haplotypes with the changes in specificity that are associated with recombination. It also explains how the proposed structure of the LRR allows different genetic rearrangements to generate a vast repertoire of resistance specificities (Fig. 4). The solvent-exposed amino acids in the LRR region may not be the only residues determining specificity;  $K_A:K_S$  analysis would not identify critical residues under selection for increased variability that are surrounded by conserved residues. However, the general aspects of the model are still valid regardless of which residues are the primary determinants of specificity.

### 1 Variation Is Generated Predominantly by Interallelic Recombination

The high level of polymorphism in ver-



## CLUSTERS OF RESISTANCE GENES

tebrate MHC genes may be generated by several mechanisms. Interallelic recombination and gene conversion at the mammalian MHC class *IB* locus generates variation substantially faster than point substitutions (Watkins et al. 1992; Parham and Ohta 1996). Phylogenetic analysis of the MHC demonstrated that polymorphic alleles, as defined by their position within the cluster, almost always formed monophyletic groups; this suggests that there has been little gene conversion between paralogs (Nei et al. 1997). Closely related mammals all have MHC class I *A*, *B*, and *C* genes; orthologs are more similar than paralogs indicating that there has been little genetic exchange between paralogs.

The model proposes that the evolution of resistance in plants is occurring primarily at the single-gene level (Fig. 4). Interallelic recombination and gene conversion could generate considerable variation and explain much of the genetic data on the instability of resistance genes. If the LRR regions of resistance-gene products have arrays of solvent-exposed  $\beta$ -strand/ $\beta$ -turn structures, there are probably multiple ligand-binding points spread along the LRR region. Recombination between alleles would result in rapid generation of new ligand-binding properties without the need for high rates of nucleotide substitution. At least some of the instability at *L* and *Rp1* observed in heteroallelic crosses could be caused by interallelic recombination. The requirement for multiple binding points would result in interallelic recombinants that had lost specificity with either combination of flanking markers, as observed with *Rp1* (Richter et al. 1995; Hulbert 1997); there is no need to invoke intergenic unequal crossing-over or gene conversion.

Unequal crossing-over between members of a tandemly repeated multigene family is not a prerequisite for variation. There is only one copy of the gene at the *L* locus in flax, yet more specificities have been identified at *L* (13) than any of the resistance clusters (Islam and Shepherd 1991; Lawrence et al. 1995). Comparisons between alleles of *L* detected variation in both the number and sequence of the LRR at the carboxy-terminal end of the predicted protein (Ellis et al. 1995, 1997). The *L2* allele has four copies of a 150-bp repeat rather than the two found in *L6*. Therefore, there can be significant variation without exchange between paralogs. Deletions within *M* and *RPP5* that resulted in losses of resistance seemed to have been caused by intragenic recombination within the LRR-encoding region (Anderson et al. 1997; Parker et al. 1997).

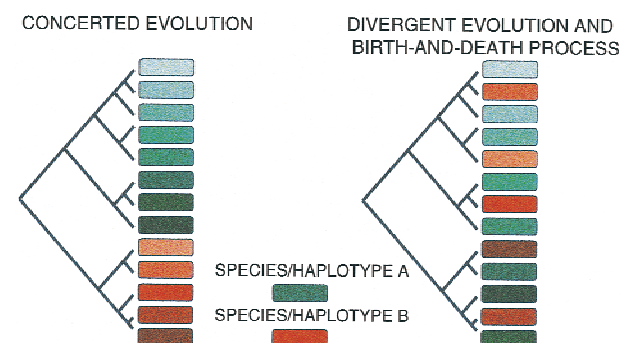
The greater sequence divergence within than between haplotypes observed at the *Pto* cluster (D.

Lavelle and R. Michelmore, unpubl.), the major cluster in lettuce (S.-S. Woo and R. Michelmore, unpubl.), and the *Cf4/9* cluster (Parniske et al. 1997), indicates that unequal crossing-over and gene conversion between paralogs is not occurring frequently enough to homogenize the sequences within a haplotype. The more rapid divergence of *Pto* pseudogene sequences supports the same conclusion (D. Lavelle and R. Michelmore, unpubl.). Therefore the pattern of sequence divergence for multiple clusters of *R* genes is more consistent with a birth-and-death process than concerted evolution (Fig. 5).

## 2. Mutations in Individual Genes are an Ultimate Source of Novel Variation

A variety of types of mutation could lead to changes in specificity. The high  $K_A:K_S$  values and ratios for regions encoding the putative solvent-exposed residues in the LRRs of *Cf4/9*, *Dm3*, *L*, *M*, *I2*, *Xa21*, and their respective homologs all indicate the significance of nonsynonymous nucleotide substitutions in these regions (Table 3). Interallelic unequal crossing-over or gene conversion could change the number of LRRs. The repeated nature of the region implies that replication slippage (Hancock 1995) is also a possibility.

The relative importance of interallelic recombination and gene conversion (1) versus de novo changes (2) in the rapid and continuous generation of new resistance specificities is impossible to evaluate at present. Recombination and gene conversion seem to be the predominant mechanisms generating allelic variation at the MHC locus (Parham and Ohta 1996). However, point mutations accumulating over long periods of time seem to be the primary source of novel polymorphism in vertebrate MHC



**Figure 5** Phylogenetic distribution of sequence variation expected for homologs under different types of evolution.

## MICHELMORE AND MEYERS

genes (Nei and Hughes 1991; Li 1997). Given that the LRR region is much larger than the binding region of the MHC protein, opportunities for interallelic recombination and gene conversion would seem to be correspondingly greater. The combination of multiple LRRS within each gene and the large number of *R* genes present in the genome provides numerous opportunities for advantageous mutations to arise. A few amino acid changes can have a large effect on specificity as shown by analysis of *Pto*, *Fen*, and related proteins (J. Rathjen, J. Chang, D. Lavelle, B. Staskawicz, and R.W. Michelson, unpubl.). Accumulation of amino acid changes over long periods of time has been proposed as the mechanism for evolution of gametophytic *S*-allele specificities (Tsai et al. 1992).

### 3. Resistance Specificities May Mature through Repeated Cycles of Selection

Evolution of new specificities may not be a one-step process. More likely, the initial interaction with a new ligand may be weak; there would then be selection for progressively tighter binding, if this results in more effective resistance. This is similar conceptually to somatic antibody maturation that occurs in T cells of vertebrates (Nossal 1992; Mueller and Jemmerson 1996). Variation in ligand binding is consistent with the variability of dominance observed for *R*-gene action and the variation in resistance phenotypes in different interactions (although there are several alternative explanations).

### 4. The Frequency of *R* Alleles Will Be Determined by the Interaction of a Variety of Forces

It seems intuitive that rare advantageous alleles of individual genes will tend to increase in the population until the frequency of virulent components of the pathogen population also increase to the point at which there is negligible selective advantage. In reality, however, the situation is likely to be much more complex. Selection will act on whole haplotypes. The rates of recombination across the cluster will influence the degree of independence of each paralog. Different paralogs within a haplotype will be under selection by the same or different pathogen populations and may confer different levels of fitness in different environments. The dynamics of individual haplotypes will become increasingly complex as the number of individual paralogs that are effective against pathogens at a particular location increases. This will be further complicated

by the degree and stability of heterogeneity in the biotic and abiotic environments.

A variety of mechanisms have been proposed to explain the maintenance of polymorphism in vertebrate MHC and immunoglobulin genes. These include linkage disequilibrium, cyclical selection, heterozygote advantage, and overdominant selection, as well as frequency-dependent selection caused by minority-dependent advantage (Nevo and Beiles 1992; Li 1997; Nei et al. 1997). Most of these selective forces are not mutually exclusive and it is still the subject of debate as to which of these mechanisms most influences the levels of MHC and immunoglobulin polymorphism. Very little is known of the population genetics of *R* haplotypes and individual *R* genes within them. This will be a productive research area as the tools are now available to investigate the influence of different evolutionary forces on *R* genes.

### 5. Unequal Crossing-Over Is Significant

Clearly some unequal crossing-over is occurring. The copy number of *R*-gene paralogs varies between haplotypes for *Cf9*, *Pto*, and *Dm3* (Anderson et al. 1996; Parniske et al. 1997; D. Lavelle and R. Michelson, unpubl.; D. Sicard, E. Nevo, and R. Michelson, unpubl.). Sequencing of multiple homologs at the *Cf4/9* and *Pto* clusters indicated the occurrence of meiotic mispairing and recombination either within or between the coding regions for *R* genes. Two near-identical paralogs are present in the *Cf2* haplotype, each of which can determine the *Cf2* specificity (Dixon et al. 1996). Sequence analysis of *Dm3* paralogs suggested at least one chimeric gene (Meyers et al. 1998a). Spontaneous mutations of *Dm3* and *Cf9* are caused by deletions, although the involvement of recombination was not demonstrated (Anderson et al. 1996; Parniske et al. 1997). The instability at *Rp1* in homoallelic crosses is associated with exchange of flanking markers and seems to be strong evidence for intra- or intergenic unequal crossing-over (Ritcher et al. 1995; Hulbert 1997). It will be interesting to confirm the underlying genetic events when this gene is cloned.

Unequal crossing-over and gene conversion are clearly important in the evolution of vertebrate MHC and immunoglobulin genes. However, although still subject to debate, these mechanisms are now not thought to be primarily responsible for generating new specificities. Instead phylogenetic analyses suggest that these families have evolved by a birth-and-death process (Nei et al. 1997). The numbers of MHC class I and II genes varies between

## CLUSTERS OF RESISTANCE GENES

haplotypes. Distantly related vertebrates have different sets of class IA genes, indicating that the haplotypes have evolved by duplication, divergence, and deletion of different paralogs. The human genome contains about 90  $V_H$ , 80  $V_K$  and 50  $V_\lambda$  genes encoding immunoglobulin components; however, the precise numbers vary between haplotypes. The numbers also vary dramatically between species. The phylogenetic trees are not correlated with the physical organization of the haplotypes; therefore each of the  $V_H$ ,  $V_K$ , and  $V_\lambda$  clusters seem to have been generated by combinations of duplication, deletion, and translocation rather than repeated duplications alone (Nei et al. 1997).

The consequences of an unequal cross-over differ depending on whether the cross-over point is located in the noncoding, intergenic regions or within the coding regions of the paralogs. Both types of exchanges result in duplication and deletion of whole genes; however, only the former generates chimeric genes with potentially new specificities or expression patterns. Sequence comparisons of paralogs of *Xa21* demonstrate the shuffling of the 5' region relative to the ORF (Song et al. 1997).

#### 6. The Stability of Resistance Will Depend on the Similarity Between Haplotypes and the Mating System

In crosses between haplotypes of close but not identical structure, there will be the potential for extensive mispairing and unequal crossing-over because of structural hybridity. *Cf4* and *Cf9* regions exhibit a fair amount of similarity. *Cf9* was stable when homozygous but unstable when heterozygous with *Cf4*; the exchanges occurred between similar sequences in intergenic regions rather than within genes (Parniske et al. 1997). Extensive regions of hemizygosity will occur between individuals with divergent haplotypes. This will result in repression of recombination as has been observed in crosses involving resistance genes introgressed from wild species (Ganal and Tanksley 1996).

Consequences of the breeding system on the structure of R haplotypes have yet to be investigated. The clusters characterized so far at the molecular level are from inbreeding species predominantly. An inbreeding mating system will tend to favor the accumulation of structural variants because they would rapidly become homozygous; this would tend to promote instability in crosses with near relatives but repress recombination in more distant crosses. Outbreeding species, such as maize, may be more unstable as duplications would tend to

be hemizygous, therefore promoting a variety of pairing possibilities. It will be interesting to compare the structure and variation of *Rp1* haplotypes in corn to those characterized from rice, tomato, lettuce, and *Arabidopsis*, all inbreeding species. Significant instability was detected when *Rp1* was homozygous and flanking markers were heterozygous. The underlying genetic events have yet to be characterized but may reflect a focusing of recombination to regions of homozygosity.

#### 7. Rapid Divergence of Intergenic Regions Will Fix Variation

Sequencing of multiple haplotypes of the *Cf4/9* region detected considerable variation in intergenic regions that appeared to have been generated by deletions. Most intergenic regions were distinct from each other. Sequencing of the *Pto* cluster revealed little sequence similarity 5' and 3' to paralogs. There were variable numbers of insertions/deletions between haplotypes. Structural variation in intergenic regions and introns would reduce meiotic mispairing (Xu et al. 1995; Dooner and Martinez-Ferez 1997) and therefore unequal crossing-over and gene conversion. Structural rearrangements restrict recombination at the *S* locus in *Brassica* species (Boyes et al. 1997). Once divergence exceeds the threshold for pairing, variants or derivatives will tend to become fixed in the haplotype.

#### 8. Duplicated Genes Will Diverge

Duplication of a sequence relieves the selective pressure on all but one copy allowing divergence. Genes with new specificities as well as pseudogenes will be generated by interallelic recombination and mutation (1 and 2 above). Genes that evolve advantageous new specificities will tend to increase in the population (4 above). Pseudogenes will tend to be lost because of infrequent unequal crossing-over (5 above).

The relative frequencies of active genes and pseudogenes have yet to be determined for R homologs. Only one pseudogene was apparent from the sequence of two haplotypes of the *Pto* cluster (D. Lavelle and R. Michelmore, unpubl.). The majority of genes sequenced from one haplotype of the major cluster in lettuce had complete ORFs and were expressed; however, some genes were clearly pseudogenes (Meyers et al. 1998b). Dysfunctional genes caused by point mutations are quite common and interspersed with functional genes in the vertebrate MHC gene clusters (Trowsdale 1995). The human

## MICHELMORE AND MEYERS

$V_H$  region contains 50 functional and 40 nonfunctional genes.

### The Role of Transposable Elements

Transposable elements of several types are major components of most plant genomes and clusters of resistance genes are no exception. Such elements could play several roles in the evolution of resistance genes. Insertion of the same element to two positions flanking a *R* gene could provide the primary duplicated sequences allowing unequal crossing-over and the initial duplication of the *R* gene sequence (Wessler et al. 1995). Insertions will tend to increase misalignment and therefore increase the chances of unequal crossing-over when hemizygous. However, when homozygous, insertions will tend to decrease the chances of misalignment and therefore contribute to the divergence of intergenic regions. Several retrotransposon-related sequences are present in unique positions between *Pto* paralogs (D. Lavelle and R. Michelson, unpubl.). Eleven different families of transposable elements were identified at the *Xa21* cluster (Song et al. 1997). Footprints left by the excision of elements can result in insertion or deletion of amino acids. A revertant of *L6* had three additional amino acids in the 5' region of the gene (Ellis et al. 1997). Transposons also provide the opportunity for changing expression patterns by adding or disrupting regulatory elements.

### The Role of Pseudogenes

Sequencing of paralogs of *Xa21*, *Cf9*, *Dm3*, and *Pto* has revealed the presence of pseudogenes (Parniske et al. 1997; Song et al. 1997; Meyers et al. 1998b; D. Lavelle and R. Michelson, unpubl.). Pseudogenes tend to evolve much faster than functional genes (Ota and Nei 1994) and therefore represent the possibility of more rapid evolution of new specificities. However, longer branch lengths for pseudogenes in phylogenetic analyses of resistance genes again indicate that neither intergenic gene conversion or unequal crossing-over play a major role in homogenizing these genes (Ota and Nei 1994; D. Lavelle and R. Michelson, unpubl.).

Given that a gene can become inactive because of mutations in regions other than those determining specificity and that recombination and gene conversion can shuffle sequences between alleles or paralogs, pseudogenes may be potential reservoirs of useful variation rather than just degenerate genes. Chicken has only one functional  $V_H$  gene

and approximately 80  $V_H$  pseudogenes in the genome. Antibody diversity is generated by somatic gene conversion between the functional gene and pseudogenes (Reynaud et al. 1989; Ota and Nei 1994). Once a functional recognition specificity has evolved, it would be advantageous to maintain it in the population, either as an expressed gene or cryptically if there is a cost to expressing numerous *R* genes.

There are little data on how many *R* genes are expressed and whether there is a cost to expressing numerous *R* genes. At least 5 of the 15 copies of *M* paralogs are expressed (Ellis et al. 1997). Also, at least half of the *Dm3* paralogs are expressed (Meyers et al. 1998b; B. Meyers, K. Shen, and R. Michelson, unpubl.). The large sequence divergence that occurs 5' to paralogs suggests that they may have different expression levels or patterns. The distribution of 5' sequences and coding regions for *Xa21* paralogs suggests the possibility of promoter shuffling, although the expression pattern of each paralog was not studied (Song et al. 1997). It seems unlikely that there is a significant cost to expressing an individual *R* gene as there are numerous such sequences in the genome and each individual message is usually rare. Most studies on the cost of resistance do not compare plants truly differing in a single resistance gene; as discussed above, all plants express numerous *R* genes, although the majority cannot be recognized experimentally. However, there may be some cost because expression of NBS-LRR genes from strong promoters is often deleterious (Mindrinos et al. 1994); although this has now been achieved for *L6* (Ellis et al. 1997). Also, there is presumably an aggregate cost to expressing many *R* genes, otherwise such genes would tend to take over the genome.

### The Role of Somatic Variation

Somatic variation is a key component in generating the broad and flexible array of specificities exhibited by the vertebrate immune system. In the absence of a circulatory system in plants, somatic variation would, on first consideration, not seem advantageous. However, the plastic and compensating growth form of plants would allow somatic sectors with new resistances to survive and contribute to the overall fitness of the plant. This would be particularly advantageous in clonally propagated species as well as long-lived species such as trees that have generation times far longer than those of potential pathogens. It will be interesting to compare variability in the specificity determining domains of



## CLUSTERS OF RESISTANCE GENES

*R* genes between different parts of long-lived individuals.

## Consequences to Breeding for Disease Resistance

There are many useful resistance specificities available. Some may be cryptic within the cultivated species; others are located in wild species. When resistance is introgressed using classical breeding procedures, it is likely that many resistance genes are being introduced simultaneously, replacing those in the recipient genotype. Apparent nonhost/race nonspecific resistance may be, at least in some cases, the result of naturally occurring pyramids of race-specific genes. Multiple resistances to *Cladosporium fulvum* were detected within the *Cf9* haplotype (Parniske et al. 1997).

A current challenge is to dissect these complex families at the molecular level and access the useful variation transgenically. This would allow the transfer of resistance across sexual incompatibility barriers. Also, when we can understand the mechanisms generating new resistance specificities, we can attempt to emulate them *ex planta* to evolve genes with altered ligand-binding characteristics and providing more effective resistance or new specificities.

## Conclusion

Plant resistance genes provide numerous opportunities for studying diversifying selection and the evolution of multigene families. Much of the current theory has been developed from studies on the evolution of the vertebrate immune system. Studies on *R* genes will provide valuable insights into the evolution of such multigene families. There are numerous *R* genes in every plant species; it is easy to generate large experimental populations and to select for rare genetic events. Also, the genetic events occurring during the coevolution of plants and pathogens can now be analyzed by comparing clusters of resistance genes using natural populations.

Clusters of *R* genes are reservoirs of variation for resistance specificities rather than rapidly evolving, dynamic groups of genes. Some genes may be expressed but not functional against current pathogen populations. Other specificities may not be expressed and therefore cryptic but accessible through a variety of genetic events. The majority of *R* genes are comprised of arrays of hypervariable potential ligand-binding sites. Interallelic recombination within these arrays may be the primary mechanism generating rapid variation in binding specificity. Some unequal crossing-over occurs and is impor-

tant; however, it is infrequent and not the primary mechanism underlying the rapid evolution of new specificities. The presence of multiple complex clusters of *R* genes, each with arrays of potential ligand-binding sites (Fig. 2), suggests how plants can generate and maintain large numbers of resistance specificities against ever-changing pathogen populations.

## Glossary

As this review draws on ideas from population genetics to structural biology, the terminology used may not be familiar to a general readership. Therefore, at the suggestion of several reviewers, we include these definitions. Further discussion of these and related terms can be found in Li (1997) or Creighton (1993).

The terminology for describing clustered multigene gene families is complex, particularly when structural rearrangements have occurred. For the purposes of this review the terms are used in the following way: **Cluster** is used to describe a genetically localized group of genes. **Gene** refers to a single member of a multigene family; each gene occupies a unique position within the cluster. **Alleles** refer to alternative forms of a single gene and not to paralogs. **Homologs** are all genes of related sequence and presumably function. **Orthologs** are genes separated by a speciation event and occupy allelic positions within the cluster. **Paralogs** are genes that have arisen by duplication events and in this review usually refer to members of a single cluster. The **haplotype** is the aggregate allelic composition across the cluster.

**Nonsynonymous nucleotide substitutions** result in changes in the amino acid sequence of the protein product. Conversely, **synonymous nucleotide substitutions** do not result in changes at the amino acid level. Their relative ratio ( $K_A:K_S$ ) is indicative of the selection pressure on the region. If there is no predominant selection on the region,  $K_A:K_S$  will be ~1. Amino acid changes tend to be deleterious to the conservation of protein structure and function; therefore most sequences are under **purifying selection**, which selects against deleterious mutations, and have a  $K_A:K_S$  ratio of <1. Occasionally, there is a selective advantage for amino acid diversity; this has been best characterized for the antigen-binding groove of the MHC protein (Hughes and Nei 1988; Li 1997). Such regions will be under **diversifying selection** and have a  $K_A:K_S$  ratio >1. The **birth-and-death** process has been proposed for the evolution of the ver-

## MICHELMORE AND MEYERS

tebrate MHC and immunoglobulin clusters that involves the expansion or contraction of the cluster (gene duplication or gene loss) by unequal crossing-over and the evolution of individual genes by diversifying selection (Nei et al. 1997).

Proteins are often modular, with distinct functions performed by different regions (**domains**) of the molecule (e.g., Kuriyan 1993; Campbell and Downing 1994). In resistance proteins, the **receptor domains** that bind to molecules originating from the pathogen directly or indirectly (**pathogen-derived ligands**) are probably distinct from the **effector domains** that are responsible for initiation of the signal transduction cascade. By inference from the known crystal structure of porcine ribonuclease inhibitor, *R* genes are hypothesized to contain a binding surface comprised of arrays of  **$\beta$ -strands** (Jones and Jones 1997; Kajava 1998); these  $\beta$ -strands have amino acids with residues that face outwards (**solvent-exposed residues**) and conserved aliphatic amino acids with hydrophobic residues buried in the interior of the protein.

## ACKNOWLEDGMENTS

We thank Brandon Gaut, Rick Kesseli, Scot Hulbert, Dina St Clair, and Stephan Abel for helpful comments on the manuscript. Our work on *Dm* genes was supported by United States Department of Agriculture National Research Initiative grant 95-37300-1571. The research on *Pto* was supported by National Science Foundation (NSF) Cooperative Agreement BIR-8920216 to Center for Engineering Plants for Resistance Against Pathogens (CEPRAP) and by CEPRAP corporate associate, Sandoz Seeds. Partial support for B.C.M. was provided by a NSF Graduate Research Fellowship.

## REFERENCES

- Anderson, P.A., P.A. Okubara, R. Arroyo-Garcia, B.C. Meyers, and R.W. Michelmore. 1996. Molecular analysis of irradiation-induced and spontaneous deletion mutants at a disease resistance locus in *Lactuca sativa*. *Mol. & Gen. Genet.* **251**: 316–325.
- Anderson, P.A., G.J. Lawrence, B.C. Morrish, M.A. Ayliffe, E.J. Finnegan, and J.G. Ellis. 1997. Inactivation of the flax rust resistance gene *M* associated with loss of a repeated unit within the leucine-rich repeat coding region. *Plant Cell* **9**: 641–651.
- Baker, B., P. Zambryski, B. Staskawicz, and S.P. Dinesh-Kumar. 1997. Signaling in plant-microbe interactions. *Science* **276**: 726–733.
- Bennetzen, J.L., S.H. Hulbert, and P.C. Lyons. 1991. Genetic fine structure analysis of a maize disease resistance gene. In *Molecular strategies of pathogens and host plants* (ed. S.P. Patil, S. Ouchi, D. Mills, and C. Vance), pp. 177–188. Springer-Verlag, Berlin, Germany.
- Bent, A.F. 1996. Plant disease resistance genes: Function meets structure. *Plant Cell* **8**: 1757–1771.
- Bent, A.F., B.N. Kunkel, D. Dahlbeck, K.L. Brown, R. Schmidt, J. Giraudat, J. Leung, and B.J. Staskawicz. 1994. *RPS2* of *Arabidopsis thaliana*: A leucine-rich repeat class of plant disease resistance genes. *Science* **265**: 1856–1860.
- Bevan, M., I. Bancroft, E. Bent, K. Love, H. Goodman, C. Dean, R. Bergkamp, W. Dirkse, M. Van Staveren, W. Stiekema et al. 1998. Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*. *Nature* **391**: 485–488.
- Botella, M.A., M.J. Coleman, D.E. Hughes, M.T. Nishimura, J.D.G. Jones, and S.C. Somerville. 1997. Map positions of 47 *Arabidopsis* sequences with sequence similarity to disease resistance genes. *Plant J.* **12**: 1197–1211.
- Boyes, D.C., M.E. Nasrallah, J. Vrebalov, and J.B. Nasrallah. 1997. The self-incompatibility (S) haplotypes of *Brassica* contain highly divergent and rearranged sequences of ancient origin. *Plant Cell* **9**: 237–247.
- Campbell, I.D. and A.K. Downing. 1994. Building protein structure and function from modular units. *Trends Biotechnol.* **12**: 168–172.
- Creighton, T. 1993. *Proteins: Structures and molecular properties*, p. 507. W.H. Freeman and Company, New York, NY.
- Crute, I.R. 1986. The genetic basis of relationships between microbial parasites and their hosts. In *Mechanisms of resistance in plant diseases* (ed. R.S.S. Fraser), pp. 80–142. Martinus Nijhoff and W. Junk, Dordrecht, The Netherlands.
- Dangl, J.L., R.A. Dietrich, and M.H. Richberg. 1996. Death don't have no mercy: Cell death programs in plant-microbe interactions. *Plant Cell* **8**: 1793–1807.
- Diers, B.W., L. Mansur, J. Imsande, and R.C. Shoemaker. 1992. Mapping *Phytophthora* resistance loci in soybean with restriction fragment length polymorphism markers. *Crop Sci.* **32**: 377–383.
- Dickinson, M.J., D.A. Jones, and J.D.G. Jones. 1993. Close linkage between the *Cf-2/Cf-5* and *Mi* resistance loci in tomato. *Mol. Plant-Microbe Interact.* **6**: 341–347.
- Dixon, M.S., D.A. Jones, J.S. Keddie, C.M. Thomas, K. Harrison, and J.D.G. Jones. 1996. The tomato *Cf-2* disease resistance locus comprises two functional genes encoding leucine-rich repeat proteins. *Cell* **84**: 451–459.
- Dooner, H.K. and I.M. Martinez-Ferez. 1997. Recombination occurs uniformly within the bronze gene, a meiotic recombination hotspot in the maize genome. *Plant Cell* **9**: 1633–1646.
- Ellis, J.G., G.J. Lawrence, E.J. Finnegan, and P.A. Anderson. 1995. Contrasting complexity of two rust resistance loci in flax. *Proc. Natl. Acad. Sci.* **92**: 4185–4188.

## CLUSTERS OF RESISTANCE GENES

- Ellis, J., G. Lawrence, M. Ayliffe, P. Anderson, N. Collins, J. Finnegan, D. Frost, J. Luck, and T. Pryor. 1997. Advances in the molecular genetic analysis of the flax-flax rust interaction. *Annu. Rev. Phytopathol.* **35**: 271–291.
- Farrara, B., T.W. Illott, and R.W. Michelmore. 1987. Genetic analysis of factors for resistance to downy mildew (*Bremia lactucae*) in lettuce (*Lactuca sativa*). *Plant Pathol.* **36**: 499–514.
- Flor, H.H. 1956. The complementary systems in flax and flax rust. *Adv. Genet.* **8**: 29–54.
- . 1965. Tests for allelism of rust resistance genes in flax. *Crop Sci.* **5**: 415–418.
- . 1971. Current status of the gene-for-gene concept. *Annu. Rev. Phytopathol.* **9**: 275–296.
- Ganal, M.W. and S.D. Tanksley. 1996. Recombination around the *Tm2a* and *Mi* genes in different crosses of *Lycopersicon peruvianum*. *Theor. Appl. Genet.* **92**: 101–108.
- Godiard, L., M.R. Grant, R.A. Dietrich, S. Kiedrowski, and J.L. Dangl. 1994. Perception and response in plant disease resistance. *Curr. Opin. Genet. Dev.* **4**: 662–671.
- Grant, M.R., L. Godiard, E. Straube, T. Ashfield, J. Lewald, A. Sattler, R.W. Innes, and J.L. Dangl. 1995. Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. *Science* **269**: 843–846.
- Hammond-Kosack, K.E. and J.D.G. Jones. 1997. Plant disease resistance genes. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**: 575–607.
- Hancock, J.M. 1995. The contribution of slippage-like processes to genome evolution. *J. Mol. Evol.* **41**: 1038–1047.
- Holub, E.B. 1997. Organization of resistance genes in *Arabidopsis*. In *The gene-for-gene relationship in host-parasite interactions* (ed. I. Crute, E. Holub, and J. Burden), pp. 5–26. CAB International, Wallingford, UK.
- Hooker, A.L. 1985. Corn and sorghum rusts. In *The cereal rusts* (ed. A.P. Roelfs and W.R. Bushnell), pp. 208–236. Academic Press, New York, NY.
- Hughes, A.L. and M. Nei. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* **335**: 167–170.
- Hulbert, S.H. 1997. Structure and evolution of the *Rp1* complex conferring rust resistance in maize. *Annu. Rev. Phytopathol.* **35**: 293–310.
- Islam, M.R. and K.W. Shepherd. 1991. Present status of genetics of rust resistance in flax. *Euphytica* **55**: 255–267.
- Jahoor, A., A. Jacobi, C.M.E. Schuller, and G. Fishbeck. 1993. Genetical and RFLP studies at the *Mla* locus conferring powdery mildew resistance in barley. *Theor. Appl. Genet.* **85**: 713–718.
- Jiang, X., M. Dreano, D.R. Buckler, S. Cheng, A. Ythier, H. Wu, W.A. Hendrickson, and N.E. Tayar. 1995. Structural predictions for the ligand-binding region of glycoprotein hormone receptors and the nature of hormone-receptor interactions. *Structure* **3**: 1341–1353.
- Jones, D.A., M.J. Dickinson, P.J. Balint-Kurti, M.S. Dixon, and J.D.G. Jones. 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. and J.D.G. Jones. 1997. The role of leucine-rich repeat proteins in plant defenses. *Adv. Bot. Res. Adv. Plant Pathol.* **24**: 89–167.
- Jorgensen, J.H. 1994. Genetics of powdery mildew resistance in barley. *Crit. Rev. Plant Sci.* **13**: 97–119.
- Kajava, A.V. 1998. Structural diversity of leucine-rich repeat proteins. *J. Mol. Biol.* **277**: 519–527.
- Kajava, A.V., G. Vassart, and S.J. Wodak. 1995. Modeling of the three-dimensional structure of proteins with the typical leucine-rich repeats. *Structure* **3**: 867–877.
- Kaloshian, I., W.H. Lange, and V.M. Williamson. 1995. An aphid-resistance locus is tightly linked to the nematode-resistance gene *Mi* in tomato. *Proc. Natl. Acad. Sci.* **92**: 622–625.
- Kanazin, V., L.F. Marek, and R.C. Shoemaker. 1996. Resistance gene analogs are conserved and clustered in soybean. *Proc. Natl. Acad. Sci.* **93**: 11746–11750.
- Kerr, E.A. and D.L. Bailey. 1964. Resistance to *Cladosporium fulvum* Cke obtained from wild species of tomato. *Can. J. Bot.* **42**: 1541–1554.
- Kesseli, R.V., I. Paran, and R.W. Michelmore. 1994. Analysis of a detailed genetic linkage map of *Lactuca sativa* (lettuce) constructed from RFLP and RAPD markers. *Genetics* **136**: 1435–1446.
- Kinoshita, T. 1993. Report of the committee on gene symbolization, nomenclature and linkage groups. *Rice Genet. Newslet.* **10**: 7–39.
- Kobe, B. and J. Deisenhofer. 1994. The leucine-rich repeat: A versatile binding motif. *Trends Biochem. Sci.* **19**: 415–421.
- . 1995. A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* **374**: 183–186.
- Kunkel, B.N. 1996. A useful weed put to work—genetic analysis of disease resistance in *Arabidopsis thaliana*. *Trends Genet.* **12**: 63–69.
- Kuriyan, J. 1993. Modular protein structures. *Curr. Sci.* **64**: 85–95.
- Lagudah, E.S., O. Moullet, and R. Appels. 1997. Map-based cloning of a gene sequence encoding a nucleotide-binding

## MICHELMORE AND MEYERS

- domain and a leucine-rich region at the *Cre3* nematode resistance locus of wheat. *Genome* **40**: 659–665.
- Lawrence, G.J., E.J. Finnegan, M.A. Ayliffe, and J.G. Ellis. 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.
- Leister, D., A. Ballvora, F. Salamini, and C. Gebhardt. 1996. A PCR-based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. *Nat. Genet.* **14**: 421–429.
- Leister, D., J. Kurth, D.A. Laurie, M. Yano, T. Sasaki, K. Devos, A. Graner, and P. Schulze-Lefert. 1998. Rapid reorganization of resistance gene homologues in cereal genomes. *Proc. Natl. Acad. Sci.* **95**: 370–375.
- Li, W.-H. 1997. *Molecular evolution*. Sinauer Associates, Sutherland, MA.
- Mackill, D. and M. Bonman. 1992. Inheritance of blast resistance in near isogenic lines of rice. *Phytopathology* **82**: 746.
- Mayo, G.M.E. and K.W. Shepherd. 1980. Studies of genes controlling specific host-pathogen interactions in flax and its rust. I. Fine structure of the *M* group in the host. *Heredity* **44**: 212–227.
- Meyers, B.C. 1998. “A molecular analysis of clusters of resistance genes in lettuce.” Ph.D. Thesis. University of California, Davis, CA.
- Meyers, B.C., D.B. Chin, K.A. Shen, S. Sivaramakrishnan, D.O. Lavelle, Z. Zhang, and R.W. Michelmore. 1998a. The major resistance gene cluster in lettuce is highly duplicated and spans several megabases. *Plant Cell* (in press).
- Meyers, B.C., K.A. Shen, P. Rohani, B. Gaut, and R.W. Michelmore. 1998b. Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. *Plant Cell* (in press).
- Michelmore, R.W. 1995. Molecular approaches to manipulation of disease resistance genes. *Annu. Rev. Phytopathol.* **15**: 393–427.
- Mindrinis, M., F. Katagiri, G.L. Yu, and F.M. Ausubel. 1994. The *A. thaliana* disease resistance gene *RPS2* encodes a protein containing a nucleotide-binding site and leucine-rich repeats. *Cell* **78**: 1089–1099.
- Mueller, C.M. and R. Jemmerson. 1996. Maturation of the antibody response to the major epitope on the self antigen mouse cytochrome *c*. *J. Immunol.* **157**: 5329–5338.
- Nei, M. and A.L. Hughes. 1991. Polymorphism and evolution of the major histocompatibility complex loci in mammals. In *Evolution at the molecular level* (ed. R.K. Selander, A.G. Clar, and T.S. Whittam), pp. 222–247. Sinauer Associates, Sunderland, MA.
- Nei, M., X. Gu, and T. Sitnikova. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc. Natl. Acad. Sci.* **94**: 7799–7806.
- Nevo, E. and A. Beiles. 1992. Selection for class II *MHC* heterozygosity by parasites in subterranean mole rats. *Experimentia* **48**: 512–515.
- Nossal, G.J.V. 1992. The molecular and cellular basis of affinity maturation of the antibody response. *Cell* **68**: 1–2.
- Ohta, T. 1991. Role of diversity selection and gene conversion in evolution of major histocompatibility complex loci. *Proc. Natl. Acad. Sci.* **88**: 6716–6720.
- Ori, N., Y. Eshed, I. Paran, G. Presting, D. Aviv, S. Tanksley, D. Zamir, and R. Fluhr. 1997. The *I2C* family from the wilt disease resistance locus *I2* belongs to the nucleotide binding, leucine-rich repeat superfamily of plant resistance genes. *Plant Cell* **9**: 521–532.
- Ota, T. and M. Nei. 1994. Divergent evolution and evolution by the birth-and-death process in the immunoglobulin  $V_H$  gene family. *Mol. Biol. Evol.* **11**: 469–482.
- Parham, P. and T. Ohta. 1996. Population biology of antigen presentation by MHC class I molecules. *Science* **272**: 67–74.
- Parker, J.E., M.J. Coleman, V. Szabo, L.N. Frost, R. Schmidt, E.A. van der Biezen, T. Moores, C. Dean, M.J. Daniels, and J.D.G. Jones. 1997. The *Arabidopsis* downy mildew resistance gene *RPP5* shares similarity to the toll and interleukin-1 receptors with *N* and *L6*. *Plant Cell* **9**: 879–894.
- Parniske, M., K.E. Hammond-Kosack, C. Golstein, C.M. Thomas, D.A. Jones, K. Harrison, B.B.H. Wulff, and J.D.G. Jones. 1997. Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato. *Cell* **91**: 821–832.
- Polzin, K.M., D.G. Lohnes, C.D. Nickell, and R.C. Shoemaker. 1994. Integration of *RPS2*, *RMD*, and *RJ2* into linkage group J of the soybean molecular map. *J. Hered.* **85**: 300–303.
- Pryor, T. and J. Ellis. 1993. The genetic complexity of fungal resistance in plants. *Adv. Plant Pathol.* **10**: 282–305.
- Reynaud, C.A., A. Dahan, V. Anquez, and J.C. Weill. 1989. Somatic hyperconversion diversifies the single  $V_H$  gene of the chicken with a high incidence in the D region. *Cell* **59**: 171–183.
- Richter, T.E., T.J. Prior, J.L. Bennetzen, and S.H. Hulbert. 1995. New rust resistance specificities associated with recombination in the *Rpl* complex in maize. *Genetics* **141**: 373–381.
- Ronald, P.C., B. Albano, R. Tabien, L. Abenes, K.S. Wu, S. McCouch, and S.D. Tanksley. 1992. Genetic and physical



## CLUSTERS OF RESISTANCE GENES

- analysis of the rice bacterial blight disease resistance locus, *Xa21*. *Mol. & Gen. Genet.* **236**: 113–120.
- Rossi, M., S.B. Goggin, I. Kaloshian, D.E. Ullman, and V.M. Williamson. 1998. The nematode resistance gene *Mi* of tomato confers resistance against the potato aphid. *Proc. Natl. Acad. Sci.* **95**: 9750–9754.
- Salmeron, J.M., G.E.D. Oldroyd, C.M.T. Rommens, S.R. Scofield, H.S. Kim, D.T. Lavelle, D. Dahlbeck, and B.J. Staskawicz. 1996. Tomato *Pf* is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the *Pto* kinase gene cluster. *Cell* **86**: 123–133.
- Sanz-Alferez, S., T.E. Richter, S.H. Hulbert, and J.L. Bennetzen. 1995. The *Rp3* disease resistance locus of maize: Mapping and characterization of introgressed alleles. *Theor. Appl. Genet.* **91**: 25–32.
- Saxena, K.M.S. and A.L. Hooker. 1968. On the structure of a gene for disease resistance in maize. *Proc. Natl. Acad. Sci.* **61**: 1300–1305.
- . 1974. A study on the structure of gene *Rp3* for rust resistance in *Zea mays*. *Can. J. Genet. Cytol.* **16**: 857–860.
- Scofield, S.R., C.M. Tobias, J.P. Rathjen, J.H. Chang, D.T. Lavelle, R.W. Michelmore, and B.J. Staskawicz. 1996. Molecular basis of gene-for-gene specificity in bacterial speck disease of tomato. *Science* **274**: 2063–2065.
- Scott, A.F., P. Heath, S. Trusko, S.H. Boyer, W. Prass, M. Goodman, J. Czelusniak, L.-Y. E. Chang, and J.L. Slightom. 1984. The sequence of the gorilla fetal globin genes: Evidence for multiple gene conversion events in human evolution. *Mol. Biol. Evol.* **1**: 371–389.
- Song, W.Y., G.L. Wang, L.L. Chen, H.S. Kim, L.Y. Pi, T. Holsten, J. Gardner, B. Wang, W.x. Zhai, L.H. Zhu, C. Fauquet, and P.C. Ronald. 1995. A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. *Science* **270**: 1804–1806.
- Song, W.Y., L.Y. Pi, G.L. Wang, J. Gardner, T. Holsten, and P.C. Ronald. 1997. Evolution of the rice *Xa21* disease resistance gene family. *Plant Cell* **9**: 1279–1287.
- Staskawicz, B.J., F.M. Ausubel, B.J. Baker, J.G. Ellis, and J.D.G. Jones. 1995. Molecular genetics of plant disease resistance. *Science* **268**: 661–667.
- Sudupak, M.A., J.L. Bennetzen, and S.H. Hulbert. 1993. Unequal exchange and meiotic instability of disease-resistance genes in the *Rp1* region of maize. *Genetics* **133**: 119–125.
- Suneson, C.A. 1960. Genetic diversity—a protection against plant diseases and insects. *Agron. J.* **52**: 319–321.
- Tang, X., R.D. Frederish, J. Zhou, D.A. Halterman, Y. Jia, and G.B. Martin. 1996. Initiation of plant disease resistance by physical interaction of AvrPto and Pto kinase. *Science* **274**: 2060–2063.
- Thomas, C.M., D.A. Jones, M.A. Parniske, K.A. Harrison, P.J. Balint-Kurti, K. Hatzixanthis, and J.D.G. Jones. 1997. Characterization of the tomato *Cf-4* gene for resistance to *Cladosporium fulvum* identifies a domain which determines recognitional specificity in *Cf-4* and *Cf-9*. *Plant Cell* **9**: 2209–2224.
- Trowsdale, J. 1995. “Both man & bird & beast”: Comparative organization of MHC genes. *Immunogenetics* **41**: 1–17.
- Tsai, D.-S., H.-S. Lee, L.C. Post, K.M. Kreiling, and T. Kao. 1992. Sequence of an S-protein of *Lycopersicon peruvianum* and comparison with other Solanaceous S-proteins. *Sex. Plant Reprod.* **5**: 256–263.
- Van der Beek, J.G., J.G. Pet, and P. Lindout. 1994. Resistance to powdery mildew (*Oidium lycopersicon*) in *Lycopersicon hirsutum* is controlled by an incomplete dominant gene *Oi-1* on chromosome 6. *Theor. Appl. Genet.* **89**: 467–473.
- Van der Biezen, E.A. and J.D.G. Jones. 1998. The NB-ARC domain: A novel signaling motif shared by plant resistance gene products and regulators of cell death in animals. *Curr. Biol.* **8**: 226–227.
- Watkins, D.I., S.N. McAdam, X. Liu, C.R. Strang, E.L. Milford, C.G. Levine, T.L. Garber, A.L. Dogon, C.I. Lord, S.H. Ghim et al. 1992. New recombinant HLA-B alleles in a tribe of South American Amerindians indicate rapid evolution of MHC class I loci. *Nature* **357**: 329–333.
- Wessler, S., T.E. Bureau, and S.E. White. 1995. LTR-retrotransposons and MITEs—important players in the evolution of plant genomes. *Curr. Opin. Genet. Dev.* **5**: 814–821.
- Witsenboer, H., R.V. Kesseli, M.G. Fortin, M. Stanghellini, and R.W. Michelmore. 1995. Sources and genetic structure of a cluster of genes for resistance to three pathogens in lettuce. *Theor. Appl. Genet.* **91**: 178–188.
- Whitham, S., S.P. Dinesh-Kumar, D. Choi, R. Hehl, C. Corr, and B. Baker. 1994. The product of the tobacco mosaic virus resistance gene *N*: Similarity to Toll and the interleukin-1 receptor. *Cell* **78**: 1101–1115.
- Xu, X., A.-P. Hsia, L. Zhang, B.J. Nikolau, and P.S. Schnable. 1995. Meiotic recombination break points resolve at high rates at the 5′ end of a maize coding sequence. *Plant Cell* **7**: 2152–2161.
- Yoshimura, A., T.W. Mew, G.S. Khush, and T. Omura. 1983. Inheritance of resistance to bacterial blight in rice cultivar Cas 209. *Phytopathology* **73**: 1409–1412.

## MICHELMORE AND MEYERS

Yu, Y.G., M.A.S. Maroof, G.R. Buss, J.M. Maughan, and S.A. Tolin. 1994. RFLP and microsatellite mapping of a gene for soybean mosaic virus resistance. *Phytopathology* **84**: 60–64.

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

Zamir, D., I. Eksteinmichelson, Y. Zakay, N. Navot, M. Zeidan, M. Sarfatti, Y. Eshed, E. Harel, T. Pleban, V. Vanoss et al. 1994. Mapping and introgression of a tomato yellow leaf curl virus tolerance gene, *TY-1*. *Theor. Appl. Genet.* **88**: 141–146.

Zhang, R., H. Cai, N. Fatima, E. Buczko, and M.L. Dufau. 1995. Functional glycosylation sites of the rat luteinizing hormone receptor required for ligand binding. *J. Biol. Chem.* **270**: 21722–21728.



## Clusters of Resistance Genes in Plants Evolve by Divergent Selection and a Birth-and-Death Process

Richard W. Michelmore and Blake C. Meyers

*Genome Res.* 1998 8: 1113-1130

Access the most recent version at doi:[10.1101/gr.8.11.1113](https://doi.org/10.1101/gr.8.11.1113)

---

### References

This article cites 90 articles, 32 of which can be accessed free at:  
<http://genome.cshlp.org/content/8/11/1113.full.html#ref-list-1>

### License

### Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

Affordable, Accurate  
Sequencing.



---

To subscribe to *Genome Research* go to:  
<https://genome.cshlp.org/subscriptions>

---