# The genetic architecture of resistance

Nevin Dale Young

Plant resistance genes (R genes), especially the nucleotide binding site leucine-rich repeat (NBS-LRR) family of sequences, have been extensively studied in terms of structural organization, sequence evolution and genome distribution. These studies indicate that NBS-LRR sequences can be split into two related groups that have distinct amino-acid motif organizations, evolutionary histories and signal transduction pathways. One NBS-LRR group, characterized by the presence of a Toll/interleukin receptor domain at the amino-terminal end, seems to be absent from the Poaceae. Phylogenetic analysis suggests that a small number of NBS-LRR sequences existed among ancient Angiosperms and that these ancestral sequences diversified after the separation into distinct taxonomic families. There are probably hundreds, perhaps thousands, of NBS-LRR sequences and other types of R gene-like sequences within a typical plant genome. These sequences frequently reside in 'mega-clusters' consisting of smaller clusters with several members each, all localized within a few million base pairs of one another. The organization of R-gene clusters highlights a tension between diversifying and conservative selection that may be relevant to gene families that are unrelated to disease resistance.

**Addresses**
Department of Plant Pathology, 495 Borlaug Hall, University of Minnesota, St. Paul, Minnesota 55108, USA; and Department of Plant Biology, 220 Biological Sciences Building, University of Minnesota, St. Paul, Minnesota 55108, USA; e-mail: neviny@tc.umn.edu

**Current Opinion in Plant Biology** 2000, **3**:285–290

1369-5266/00/$ – see front matter
© 2000 Elsevier Science Ltd. All rights reserved.

**Abbreviations**

| | |
|---|---|
| **BAC** | bacterial artificial chromosome |
| **CC** | coiled-coil |
| **IL** | interleukin |
| **kb** | kilobases |
| **LRR** | leucine-rich repeat |
| **Mb** | Megabases |
| **NBS** | nucleotide binding site |
| **PCR** | polymerase chain reaction |
| **R** | resistance |
| **TIR** | Toll/interleukin receptor |

## Introduction

Disease resistance (R) genes comprise a large and diverse group of related sequences in plant genomes. Most plant R genes seem to be members of an ancient gene family that encode nucleotide-binding proteins [1••,2••]. Because of their distinctive domain structure, these proteins are known as 'nucleotide binding site-leucine rich repeats' (NBS-LRRs). Research on the genes encoding NBS-LRRs and other R proteins sheds light on the global genome organization, sequence variability and evolutionary history of R genes. The results of this research provide a pivotal foundation for understanding the generation of diversity, molecular evolution, mechanisms of recognition and modes of signal transduction involving R proteins. The discoveries arise from large-scale genome sequencing, re-sequencing of R gene clusters, the use of degenerate PCR (polymerase chain reaction) primers to harvest R-gene candidates, comparative genomics and phylogenetic analysis. Increased understanding of R genes also reflects the growing power of computational biology to describe and model genome structure.

In addition to NBS-LRRs, other types of plant R proteins have been described. For example, several plant proteins consisting of an LRR bound to a transmembrane domain have been shown to be encoded by functional R genes, including the *Hcr9* family of tomato [3]. The *Hcr9* family is known to provide resistance to different races of *Cladosporium fulvum*. The tomato gene *Pto* is a serine-threonine kinase that confers resistance to *Pseudomonas syringae* [4], the rice gene *Xa21* combines LRR and kinase domains to confer resistance to *Xanthomonas oryzae* [5] and the sugar beet gene *Hs1pro-1* contains a modified LRR structure and confers resistance to *Heterodera schachtii* [6]. Nevertheless, the NBS-LRR family is the largest group of known R-gene products, including at least 15 known R genes from six plant species [1••]. NBS-LRRs have been the focus of intense research during the past few years and the fruits of this research are the primary subject of this review. For excellent previous reviews on plant R genes, see [7–9].

## Structural organization of NBS-LRR sequences

NBS-LRR proteins possess a putative nucleotide binding domain that is probably involved in signal transduction [8]. This domain is composed of several short amino-acid motifs that are highly conserved among family members and that are interspersed among other largely divergent sequences [1••,2••]. The conserved motifs have been extensively characterized and are described in further detail below. Plant NBS domains show sequence similarity to nematode *CED-4* and mammalian *Apaf-1*, which have been implicated in protease-mediated apoptosis [10]. *Apaf-1* has also been shown to form oligomers [11], which may be relevant in the function of plant R-gene NBS domains.

Carboxy-terminal to the NBS, R genes usually contain an LRR domain that is thought to be involved in ligand binding and pathogen recognition. Other types of R genes, including the *Hcr9* family, *Cf2/5* and *Xa21*, also encode LRR domains. LRRs consist of repeated imperfect amino-acid segments that fold into solvent exposed β-strand β-turn structures [12]. In studies on NBS-LRRs, Meyers *et al.* [13] systematically compared the LRR regions of the lettuce *Dm3*, tomato *I2C*, tomato *Mi*, rice *Xa21* and flax *L/M* gene families and found an alternating pattern of conservation and hypervariability. The variability was highest for codons (x) positioned around the

two conserved aliphatic amino acids in the LRR consensus xx(a)x(a)xx (where a is the conserved aliphatic amino acid). Based on the ratio of synonymous to non-synonymous amino-acid substitutions, the results indicate that these sites have experienced diversifying selection. Nonetheless, detailed studies on *RPS2* [14] and *Rpm1* [15] of *Arabidopsis* suggest that variation at disease resistance loci can also be explained by other mechanisms. The actual number of LRR repeats can differ among family members. Among sequences at the *Cf2/5* locus of tomato, the number of LRR repeats ranges from 25 to 38 [16]. Finally, in studies of the *L* locus of flax, experimentally-derived chimeric proteins exhibited the pathogen specificity of either the donor LRR domain or entirely novel specificities [17•]. Together, these observations provide growing support for the model in which LRR domains undergo high rates of change that are crucial in the evolution of pathogen recognition. LRR domains may also play a role in signal transduction. Warren *et al.* [18] found that a single amino-acid substitution within the LRR domain of *Arabidopsis RPS5* partially compromises the functions of several R-gene responses. An extensive review of LRRs and their possible roles in plant resistance can be found in [12].

Another striking feature of some NBS-LRR proteins is the presence of an amino-terminal domain showing homology to both Toll of *Drosophila* and interleukin-receptor-like proteins of mammals [19]. This domain is generally referred to as the TIR (Toll/interleukin receptor) domain. Toll, IL-1R (i.e. interleukin-1R) and related proteins have been shown to be involved in non-specific cellular immunity in animals. By analogy to these animal proteins, plant TIRs are thought to function in signal transduction [8]. However, recent work suggests that TIRs may also be involved in pathogen recognition. In a study of 13 alleles of the flax resistance gene *L*, variation in the TIR domain was associated with changes in pathogen recognition [17•]. The authors sequenced 13 alleles at the *L* locus, each conferring a distinct specificity for a different flax rust isolate. Two alleles with differing specificities were found to possess changes only in their TIR domains. Thus, it appears that both the TIR and LRR domains play a role in pathogen recognition. Indeed, a bioinformatic survey of known R-genes and predicted TIR-NBS-LRR sequences of *Arabidopsis* indicated there are at least eight different configurations of conserved motifs within the TIR domain [1••].

Previously, many NBS-LRRs lacking an amino-terminal TIR were thought to possess a leucine-zipper motif [20,21]. However, recent analysis of a wide array of NBS-LRRs indicates the most striking feature may be a coiled-coiled (CC) structure in place of the TIR [2••]. CCs are bundles of two to five helices that have a distinctive packing of amino-acid side chains at the helix–helix interface [22]. The CC structure typically exhibits a seven-residue-repeat organization with the hydrophobic side chains of two of the amino acids forming an interface for interactions between coils. Leucine zippers are members of this broader class of structural elements. Pan *et al.* [2••] have referred to NBS-LRR sequences bound to a TIR domain as 'Group I' and those without a TIR as 'Group II', whereas Meyers *et al.* [1••] refer to these groups as 'TIR' and 'non-TIR'.

TIR and non-TIR NBS-LRRs can also be distinguished by the amino-acid motifs found within the NBS domain itself. Nearly all plant NBS domains are characterized by several highly conserved amino-acid motifs, including the well-described P-loop, Kin-1a and 'GLPL' sites [1••,2••]. However, recent analysis demonstrates that most TIR-NBS-LRRs also contain a stretch of conserved amino acids just after the P-loop with the consensus sequence LQKKLLSKLL (using single-letter code for amino acids), as well as a second amino acid-motif preceding the LRR domain (FLHIACFF). (For both of the motifs described in this paragraph, sequences indicate the most frequent amino-acid residue at each position. Biochemically similar amino acids can also be present.) Both are absent from non-TIR NBS-LRRs. In contrast, non-TIR NBS-LRRs typically contain a distinctive amino-acid motif near the P-loop (FDLxAWVCVSQxF) and another motif near the carboxy-terminus of the NBS domain (CFLYCALFP), both of which are absent from TIR-containing NBS-LRRs. A single residue in the highly conserved motif within the NBS known as kinase-2 (LLVLDDVW/D) can be used to predict the presence of the TIR domain with 95% accuracy; a tryptophan (W) residue is found in non-TIR proteins whereas an aspartic acid (D) residue is found in TIR-containing proteins. Overall, these motifs are so diagnostic that it has been possible to develop degenerate primers that specifically amplify either one of the two groups of NBS-LRRs ([2••]; S Peñuela, ND Young, unpublished data). These observations are especially important for researchers who hope to 'harvest' R-gene sequences from plant genomes using degenerate primers. First, it should now be possible to design group-specific primers that amplify specific subsets of NBS sequences. Second, researchers should have a road map for designing combinations of primers that are most likely to uncover the widest array of resistance-gene candidates.

Structural differences between TIR and non-TIR NBS-LRR sequences appear to have functional significance as well. Two different *Arabidopsis* mutations, *eds1* [23] and *ndr1* [22], have been shown to eliminate R responses to multiple pathogens. *NDR1* encodes a putative transmembrane protein of unknown function [24], whereas *EDS1* encodes a putative lipase [25]. The mutations are thought to affect the signal transduction pathway downstream of R-gene mediated recognition of pathogens. Further analysis of these mutants indicates that TIR-NBS-LRR sequences, including *RPP5*, operate through an *EDS1*-dependent pathway, whereas many non-TIR NBS-LRR sequences, including *RPS2* and *RPM1*, operate through a *NDR1*-pathway [26]. Conceivably, the amino-terminal TIR or CC domains and/or related NBS motifs could play a role in the bifurcation of the signal transduction pathway.

## Phylogenetic analysis of NBS-LRR sequences

Phylogenetic analysis of NBS-LRR sequences supports the division into TIR and non-TIR groups [1••,2••]. One noteworthy result of these studies has been the observation that TIR NBS-LRRs are undetectable and probably absent in the Poaceae. Efforts to uncover TIR-containing sequences in the Poaceae in large public databases, as well as even larger proprietary databases, have uniformly failed. In contrast, non-TIR NBS-LRR sequences are found in all of the Angiosperm species tested. Targeted PCR amplification experiments add further support to these observations [2••]. Degenerate PCR primers specific for either TIR or non-TIR sequences successfully amplified both types of products with tomato genomic DNA as template, but only non-TIR products when wheat was used as template. To explain these observations, Pan *et al*. [2••] noted that a TIR-containing NBS-LRR sequence has been reported in *Pinus* (Genbank accession number AF038649) as well as the fact that TIRs are found in animals. This led them to propose a model in which the common ancestor of Angiosperms and Gymnosperms contained both types of NBS-LRR sequences with the branch leading to modern grasses losing the TIR class of NBS-LRR sequences after divergence. It will be especially interesting to learn whether other monocots lack TIR sequences. Indeed, it will also be valuable to examine other Gymnosperms and more ancient taxa of green plants to discover what sort of NBS-LRRs, if any, exist in the ancestors of Angiosperms. Preliminary results suggest that NBS-LRR sequences do exist among primitive land plants [27].

Phylogenetic studies have also shed light on diversity within the NBS-LRR family. Phylogenies of NBS-LRR sequences tend to have long branch lengths and closely clustered nodes, indicating ancient divergence into separate lineages followed by more recent diversification [1••,2••]. Phylogenies of TIR-NBS-LRR sequences contain several distinct subgroups of sequences, reflecting recent diversification within individual species or closely related plant taxa. Several subgroups contain sequences from closely related species that are present multiple times within a single species. This indicates that some TIR-NBS-LRR sequences have diverged both prior to and since speciation. Trees of non-TIR sequences are composed almost exclusively of species- or family-specific clades, though some branches containing sequences from multiple taxa do exist. Still, monocot sequences are not monophyletic with respect to dicot sequences, suggesting that the ancient ancestor of moncots and dicots contained multiple non-TIR NBS-LRR sequences that have since diverged [1••,2••]. Finally, nearly every branch of both TIR and non-TIR trees contains at least one confirmed R gene, suggesting that most NBS-containing sequences are similar to known R genes and may therefore encode functional R proteins. This is especially significant because so many of the NBS-LRR sequences reported to date have been isolated by PCR amplification with degenerate primers and no direct connection to actual resistance phenotypes.

In a detailed phylogenetic analysis of tomato and other Solanaceous NBS-LRR sequences, Pan *et al*. [28] observed that sequences from different Solanaceous species are well distributed among branches of the tomato phylogenetic tree. This suggests that these NBS-LRRs probably arose from common ancestors that existed before speciation within the Solanaceae. Potato and tomato sequences form tight clusters on the phylogenetic tree, and some tomato NBS-LRRs are more closely related to potato than other tomato-derived sequences. Potentially these NBS-LRRs represent sequence orthologs or sequences derived from common ancestors. Even after dozens of *Arabidopsis* NBS-LRRs were added to the phylogenetic tree, all branches remained plant family specific. From this, Pan *et al*. infer that major gene duplication events occurred during dicot divergence into various taxa, followed by recent radiation from common ancestors.

## Genomic architecture of R-gene sequences

The NBS-LRR family of sequences is one of the largest known in plants. Recent estimates for the number of NBS-LRR sequences in *Arabidopsis* range from 200–300 [1••,2••]. In rice, the estimates are even higher (though far less certain) with values as high as 1500. Even if present at the lower frequencies estimated for *Arabidopsis*, NBS-LRR sequences probably account for 1–2% of the total coding capacity of the genome (assuming a total of 21,000 genes [29]). On chromosome IV alone, resistance-related LRR-containing sequences (primarily NBS-LRR and *Hcr9*-like sequences) account for nearly 3.7% of all gene sequences [30•]. Based on genomic-sequencing analysis of the NBS-LRR sequences of *Arabidopsis*, TIR-containing sequences outnumber non-TIR sequences by roughly three to one [1••].

Most NBS-LRR and other R-gene-like sequences reside in large, extended arrays. Some of these 'mega-clusters' can be huge, spanning millions of base pairs and consisting of dozens of R-gene sequences. The physical structure of these clusters is thought to be involved in both the generation and maintenance of R-gene diversity. In *Arabidopsis*, for example, several NBS-LRR sequences on chromosome IV are co-localized with the phenotypically defined R-gene cluster known as MRC-H [31]. This mega-cluster, which has been partially described in [1••,29,32], includes two smaller clusters that each have seven or eight NBS-LRR sequences [30•]. These two clusters are located approximately 1 Megabase (Mb) apart near the center of the chromosome. One of the clusters contains *RPP5* plus seven other NBS-LRR genes over a stretch of 90 kilobases (kb). Relative to the remainder of the *Arabidopsis* genome, this cluster exhibits extremely high levels of intraspecific polymorphism and many of the NBS-LRR sequences are pseudogenes [32]. The two clusters are located near four smaller NBS-LRR clusters that together contain 11 additional sequences, with all six clusters located within a 4.6 Mb stretch [33]. In a second mega-cluster on *Arabidopsis* chromosome V, there are more than 14 NBS-LRR

sequences within 300 kb co-localized with MRC-J [34]. Of these sequences, 12 are TIR-containing sequences in two tandem arrays along with two isolated non-TIR NBS-LRRs. Nearby, there are eight additional NBS-LRR sequences within 300 kb. In fact, there are a total of 30 NBS-LRRs in the 4.5 Mb region encompassing these sequences on chromosome V. At least some of the sequences appear to encode TIR domains not associated with NBS-LRRs.

Molecular studies based on long-range physical mapping and DNA sequencing suggest that R genes in other plant species are also organized in large clusters. The *M* locus of flax consists of 15 or more gene family members spread over a distance of less than 1 Mb [35], the *Xa21* locus of rice consists of eight or more sequences spanning 230 kb [36] and the *Cf4/9* locus of tomato contains five closely related members spanning 35 kb [37]. In a detailed study of the *Dm3* cluster of lettuce, at least 24 non-TIR NBS-LRR sequences were found to span approximately 3.5 Mb [38]. To characterize this region, the authors used a combination of high resolution genetic mapping, deletion break-point lines and bacterial artificial chromosome (BAC) clones — an indication of the scale of work required to dissect the structure of extended genomic regions in non-model organisms. The data suggest that the spacing between NBS-LRRs in the *Dm3* cluster is at least 150 kb, and sequence sampling throughout the region indicates few, if any, intervening functional genes apart from NBS-LRRs. Of these *Dm3*-like sequences, at least eight appear to be expressed.

## R-gene clusters and the evolution of specificity

Equipped with detailed knowledge of the genomic organization of NBS-LRR sequences, it is possible to ask whether physical genomic distance is related to phylogenetic distance; in other words, are genes that are close together on the chromosome also closely related in evolution? Examined at the fine structure level, physical and phylogenetic distance do not strictly correspond at the *Dm3* cluster [38]. Sequences that are adjacent on the genome frequently lie on separate branches of the phylogenetic tree. For example, two physically adjacent NBS-LRR sequences in the *Dm3* cluster (*RGC2I* and *RGC2B*, the *Dm3* locus itself) lie on distinct and strongly supported branches on the phylogenetic tree with only 71% amino-acid similarity. In contrast, four other NBS-LRR sequences lie on the chromosome between *RGC2B* and the sequence to which it is most closely related evolutionarily, *RGC2S*. In other cases, NBS-LRR gene clusters have even been shown to consist of members from distinct sub-families. At the *Mla* cluster of barley, for example, at least three different sub-families of NBS-LRRs are found within a 240 kb region and the amino-acid similarity among the sub-families is as low as 33% [39•].

One possible explanation for the existence of of R-gene clusters composed of different families of sequences could be interlocus exchange between physically separated clusters. This has been shown to have occurred in the *Hcr9* gene family of tomato. *Hcr9* genes reside in three distinct clusters spanning 20 centimorgans on chromosome I [40]. Sequence analysis indicates that each cluster is composed of sequences that are more closely related to one another than to those of neighboring clusters, an indication that *Hcr9* duplication and movement preceded diversification within clusters. However, one exceptional member exhibits sequence features that are specific for *Hcr9* sequences at separate clusters, suggesting a recent interlocus recombination event [41••]. The possibility that R-gene sequences translocate from one genomic location to another goes directly to the question of how R-gene diversity might be generated and maintained within large, extended clusters. Reservoirs of non-identical R-gene sequences coming from distant genomic locations would certainly provide the starting material for creating novel specificities.

Nevertheless, frequent unequal crossing-over and gene conversion events would inevitably lead to sequence homogenization and concerted evolution within clusters. Comparisons between R haplotypes (defined in [42] as the aggregate allelic composition across a gene cluster) reveal that orthologs (sequences separated by speciation and occupying allelic positions within a gene cluster) are generally more similar than paralogs (duplicated sequences within a gene cluster). Michelmore and Meyers [42] interpret these observations as evidence for a 'birth and death' model of plant R-gene evolution. Under this model, intergenic unequal crossing-over and gene conversion are important in creating new members, but novel specificities derive primarily from divergent selection acting on pathogen recognition regions within R genes. Given the results with the *Hcr9* gene family, this model must be expanded to include the possibility that interlocus recombination between R-gene clusters also occurs. In this way, novel sequence combinations could be introduced by exchange between physically separate loci. Following such events, unequal crossing-over and gene conversion would initiate novel lineages of sequence evolution.

## Is the genetic architecture of R genes special?

It is clear that NBS-LRR and other R genes are organized in large, extended clusters in the genome, but how about other plant gene families? The recent publication of complete genome sequences for chromosomes II and IV of *Arabidopsis* reveals that gene clusters and genome duplications are quite commonplace — even in a genome that is as structurally 'simple' as *Arabidopsis* [28,43•,44]. Nearly 12% of the genes and predicted genes on chromosome IV exist in gene clusters, primarily as tandem duplications. Clusters with larger numbers of related sequences are also common, including one with 15 contiguous receptor-kinase-like proteins (of course, the possibility that these genes might be involved in resistance can not be ruled out). Additional large clusters consist of genes encoding cytochrome-P450-like proteins as well as proteins involved in a wide variety of metabolic pathways. On chromosome II, a single BAC clone (F16P2) contains repeats of 12 putative tropinone

reductase genes plus another array with seven glutathione S-transferase genes. Is it possible that new substrate specificities might be generated through this type of genome organization as tandem gene clusters, just like R genes? Certainly it makes sense that the balancing forces of conservative and diversifying selection should play a role in gene families other than R genes? If so, detailed studies of the genome architecture and molecular evolution of plant R genes could have important implications beyond the field of plant–microbe interactions.

## Conclusions

Over the past few years, the sequence organization, genome distribution and evolutionary history of plant R genes, especially NBS-LRRs, have become much better understood. The ancient nature of NBS-LRR sequences, their separation into distinct lineages and more recent diversification helps to explain the observed sequence diversity and structural features of this gene family. At a genomic level, extensive gene clusters are a striking property of most R genes that is probably related to a balance between creating new specificites and conserving old ones. The possibility of exchanges between clusters magnifies the opportunities for generating novel specificities. Future research must integrate our growing knowledge of R-gene sequence diversity, pathogen recognition and genome organization. Extensive genome sequencing, especially re-sequencing of R-gene clusters, will provide valuable data. At the same time, new bioinformatic tools and coordinated efforts in structural and functional genomics will be essential.

## Update

Recent work is beginning to unravel the molecular interactions of NBS-LRR proteins. Using a yeast two-hybrid system, Van der Biezen and co-workers discovered that the *Arabidopsis* TIR-NBS-LRR gene, *RPP5*, interacts with *At-RSH1* [45]. This plant protein shows strong homology to *RelA* and *SpoT* of *Escherichia coli*, proteins that determine the levels of guanosine tetraphosphate and guanosine pentaphosphate. Thus, it is possible that these molecules could have a role as second messengers in plants through which R-genes such as *RPP5* communicate the recognition of pathogens.

## Acknowledgements

I wish to thank Steve Canon, Allan Dickerman, Blake Meyers and Deborah Samac for valuable advice and insightful suggestions, as well as Doug Cook, Robert Fluhr and James Groth for helpful comments. I also thank Silvia Peñuela for providing unpublished data.

## References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest


1. Meyers BC, Dickermann AW, Michelmore RW, Sivaramakrishnan S,
•• Sobral BW, Young ND: **Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide binding superfamily.** *Plant J* 1999, **20**:317-332.
This paper is a comprehensive examination of motif organization, phylogeny and genomic evolution among NBS-LRR proteins. Descriptions of TIR- versus non-TIR-specific motifs, an extensive phylogenetic tree of NBS-LRR evolution and the distribution of NBS-LRR sequences throughout the *Arabidopsis* genome are especially noteworthy. The study also provides evidence that TIR-containing sequences are absent from the grasses.

2. Pan Q, Wendel J, Fluhr R: **Divergent evolution of plant NBS-LRR**
•• **resistance gene homologues in dicot and cereal genomes.** *J Mol Evol* 2000, **50**:203-213.
This paper is another comprehensive overview of plant NBS-LRR structure and phylogeny. In addition to identifying group-specific motifs and examining the apparent absence of TIR-NBS-LRR sequences in grasses, the study lays out the evidence for the presence of CC structures in the amino-terminal region of many NBS-LRR proteins.

3. Jones DA, Thomas CM, Hammond-Kosack KE, Balint-Kurti PJ, Jones JDG: **Isolation of the tomato *Cf-9* gene for resistance to *Cladosporium fulvum* by transposon tagging.** *Science* 1994, **266**:789-793.

4. Martin GB, Brommonschenkel SH, Chunwongse J, Frary A, Ganal MW, Spivey R, Wu T, Earle ED, Tanksley SD: **Map-based cloning of a protein kinase gene conferring disease resistance in tomato.** *Science* 1993, **262**:1432-1436.

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

6. Cai D, Kleine M, Kifle S, Harloff H-J, Sandal NN, Marcker KA, Klein-Lankhorst RM, Salentijn EMJ, Lange W, Stiekema WJ *et al.*: **Positional cloning of a gene for nematode resistance in sugar beet.** *Science* 1997, **275**:832-834.

7. Bent AF: **Plant disease resistance genes: function meets structure.** *Plant Cell* 1996, **8**:1757-1771.

8. Ellis J, Jones D: **Structure and function of proteins controlling strain-specific pathogen resistance in plants.** *Curr Opin Plant Biol* 1998, **1**:288-293.

9. Ronald PC: **Resistance gene evolution.** *Curr Opin Plant Biol* 1998, **1**:294-298.

10. Li P, Nijhawan D, Budihardjo I, Srinvasula SM, Ahmad M, Alnemri ES, Wang X: **Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade.** *Cell* 1997, **91**:479-487.

11. Srinivasula SM, Ahmad M, Fernandes-Alnemri T, Alnemri ES: **Autoactivation of procaspase-9 by Apaf-1-mediated oligomerization.** *Mol Cell* 1998, **1**:949-957.

12. Jones DA, Jones JDG: **The role of leucine-rich repeat proteins in plant defences.** *Adv Bot Res* 1997, **24**:90-167.

13. Meyers BC, Shen KA, Rohani P, Gaut BS, Michelmore RW: **Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection.** *Plant Cell* 1998, **11**:1833-1846.

14. Caicedo AL, Schaal BA, Kunkel BN: **Diversity and molecular evolution of the *RPS2* resistance gene in *Arabidopsis thaliana*.** *Proc Natl Acad Sci USA* 1999, **96**:302-306.

15. Stahl EA, Dwyer G, Mauricio R, Kreitman M, Bergelson J: **Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*.** *Nature* 1999, **400**:667-671.

16. Dixon MS, Hatzixanthis K, Jones DA, Harrison K, Jones JDG: **The tomato *Cf-5* disease resistance gene and six homologs show pronounced allelic variation in leucine-rich repeat copy number.** *Plant Cell* 1998, **10**:1915-1925.

17. Ellis JG, Lawrence GJ, Luck JE, Dodds PN: **Identification of regions**
• **in alleles of the flax rust resistance gene *L* that determine differences in gene-for-gene specificity.** *Plant Cell* 1999, **11**:495-506.
This paper presents a sequence comparison of 13 alleles at the *L* locus of flax. Comparisons of different alleles, as well as transgenic plants with chimeric R genes, confirm the importance of the LRR domain and provide evidence that both the TIR and the LRR domains are involved in pathogen specificity.

18. Warren RF, Henk A, Mowery P, Holub E, Innes RW: **A mutation within the leucine-rich repeat domain of the *Arabidopsis* disease resistance gene *RPS5* partially suppresses multiple bacterial and downy mildew resistance genes.** *Plant Cell* 1998, **10**:1439-1452.

19. Hoffmann JA, Kafatos FC, Janeway CA, Ezkowitz RAB: **Phylogenetic perspectives in innate immunity.** *Science* 1999, **284**:1313-1317.

20. Grant MR, Godiard L, Straube E, Ashfield T, Lewald J, Sattler A, Innes RW, Dangl JL: **Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance.** *Science* 1995, **269**:843-846.

21. Bent A, Kunkel BN, Dahlbeck D, Brown KL, Schmidt R, Giraudat J, Leung J, Staskawicz BJ: ***RPS2* of *Arabidopsis thaliana*: a leucine-rich repeat class of plant disease resistance genes.** *Science* 1994, **265**:1856-1860.

22. Lupas A: **Coiled coils: new structures and new functions.** *Trends Biochem Sci* 1996, **21**:375-382.

23. Parker JE, Holub EB, Frost LN, Falk A, Gunn ND, Daniels MJ: **Characterization of *eds1*, a mutation in *Arabidopsis* suppressing resistance to *Peronospora parasitica* specified by several different *RPP* genes.** *Plant Cell* 1996, **8**:2033-2046.

24. Century KS, Shapiro AD, Repetti PP, Dahlbeck D, Holub E, Staskawicz BJ: ***NDR1*, a pathogen-induced component required for *Arabidopsis* disease resistance.** *Science* 1997, **278**:1963-1965.

25. Falk A, Feyes B, Frost LN, Jones JDG, Daniels MJ, Parker JE: ***EDS1*, an essential component of R gene-mediated disease resistance in *Arabidopsis* has homology to eukaryotic lipases.** *Proc Natl Acad Sci USA* 1999, **96**:3292-3297.

26. Aarts N, Metz M, Holub E, Staskawicz BJ, Daniels MJ, Parker JE: **Different requirements for *EDS1* and NDR1 by disease resistance genes define at least two R gene-mediated signaling pathways in *Arabidopsis*.** *Proc Natl Acad Sci USA* 1998, **95**:10306-10311.

27. Goodwin SB: **High evolutionary conservation of resistance gene analogs from diverse plant hosts.** In *Plant & Animal Genome VII*. 1999. San Diego: Scherago International, Inc.:189.

28. Pan Q, Liu Y-S, Budai-Hadrian O, Sela M, Carmel-Goren L, Zamir D, Fluhr R: **Comparative genetics of NBS-LRR resistance gene homologues in the genomes of two dicotyledons: tomato and *Arabidopsis*.** *Genetics* 2000, in press.

29. European Union *Arabidopsis* Genome Project: **Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of *Arabidopsis thaliana*.** *Nature* 1998, **391**:485-488.

30. European Union *Arabidopsis* Genome Sequencing Consortium:
• **Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.** *Nature* 1999, **402**:769-777.

One of two recent landmark studies in which the entire sequence of a plant chromosome is described. Of special interest in the context of R genes is the revelation that gene duplications and extensive gene clusters are relatively common in *Arabidopsis*.

31. Holub E: **Organization of resistance genes in *Arabidopsis*.** In *The Gene-for-Gene Relationship in Plant–Parasite Interactions*. Edited by Crute IR, Holub EB, Burdon JJ. Farnham, UK: CAB International; 1997:5-26.

32. Noel L, Moores TL, van der Biezen EA, Parniske M, Daniels MJ, Parker JE, Jones JDG: **Pronounced intraspecific haplotype divergence at the *RPP5* complex disease resistance locus of *Arabidopsis*.** *Plant Cell* 1999, **11**:2099-2111.

33. Mewes HW, Heumann K, Kaps A, Mayer K, Pfeiffer F, Stocker S, Frishman D: **MIPS: a database for genomes and proteins sequences.** *Nucleic Acid Res* 1999, **27**:44-48.

34. Carnegie Institute of Washington, National Center for Genome Resources: **The *Arabidopsis* Information Resource.** URL http://www.arabidopsis.org/

35. Anderson PA, Lawrence GJ, Morrish BC, Ayliffe MA, Finnegan EJ, Ellis JG: **Inactivation of the flax rust resistance gene *M* associated with loss of a repeated unit within the leucine-rich repeat coding region.** *Plant Cell* 1997, **9**:641-651.

36. Williams CE, Wang B, Holsten TE, Scambray J, Dasilva FDG, Ronald PC: **Markers for selection of the rice *Xa21* disease resistance gene.** *Theor Appl Genet* 1996, **93**:1119-1122.

37. Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA, Harrison K, Wulff BBH, Jones JDG: **Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the *Cf-4/9* locus of tomato.** *Cell* 1997, **91**:821-832.

38. Meyers BC, Chin DB, Shen KA, Sivaramakrishnan S, Lavelle DO, Zhang Z, Michelmore RW: **The major resistance gene cluster in lettuce is highly duplicated and spans several megabases.** *Plant Cell* 1998, **10**:1817-1832.

39. Wei F, Gobelman-Werner K, Morroll SM, Kurth J, Mao L, Wing R,
• Leister D, Schulze-Lefert P, Wise RP: **The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley.** *Genetics* 1999, **153**:1929-1948.

The structure of the *Mla* gene cluster of barley is described. To examine these sequences, a complete physical map consisting of BAC and yeast artificial chromosome clones spanning more than 1 Mb is constructed and a segment with 11 NBS-LRR sequences is pinpointed. Interestingly, rates of recombination were found to be ten-fold lower in this region than in the genome at large.

40. Parniske M, Wulff BBH, Bonnema G, Thomas CM, Jones DA, Jones JDG: **Homologues of the *Cf-9* disease resistance gene (*Hcr9s*) are present at multiple loci on the short arm of tomato chromosome 1.** *Mol Plant Microbe Interact* 1999, **12**:93-102.

41. Parniske M, Jones JDG: **Recombination between diverged clusters**
•• **of the tomato *Cf-9* plant disease resistance gene family.** *Proc Natl Acad Sci USA* 1999, **96**:5850-5855.

This is a fascinating study of three *Hcr9* clusters in tomato in which sequence analysis of family members suggests that recombination has taken place between clusters. This observation could have profound implications for the genome organization and evolution of R genes, as well as other dispersed gene families.

42. Michelmore RW, Meyers BC: **Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process.** *Genome Res* 1998, **8**:1113-1130.

43. Lin X, Kaul S, Rounsley S, Shea TP, Benito MI, Town CD, Fujii CY,
• Mason T, Bowman CL, Barnstead M *et al.*: **Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.** *Nature* 1999, **402**:761-768.

One of two recent landmark studies in which the entire sequence of a plant chromosome is described. Of special interest in the context of R genes is the revelation that gene duplications and extensive gene clusters are relatively common in *Arabidopsis*.

44. Martienssen R, Irish V: **Copying out our ABCs — the role of gene redundancy in interpreting genetic hierarchies.** *Trends Genet* 1999, **15**:435-437.

45. Van der Biezen EA, Sun J, Coleman MJ, Bibb MJ, Jones JDG: ***Arabidopsis* RelA/SpoT homologs implicate (p)ppGpp in plant signaling.** *Proc Natl Acad Sci USA* 2000, **97**:3747-3752.