MOLECULAR PLANT PATHOLOGY (2006) **7**(5), 437–448 DOI: 10.1111/J.1364-3703.2006.00342.X

Review

Recent insights into *R* gene evolution

JOHN M. MCDOWELL* AND STACEY A. SIMON

Department of Plant Pathology, Physiology, and Weed Science, and Fralin Biotechnology Center, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061-0346, USA

SUMMARY

Plants are under strong evolutionary pressure to maintain surveillance against pathogens. Resistance (*R*) gene-dependent recognition of pathogen avirulence (*Avr*) determinants plays a major role in plant defence. Here we highlight recent insights into the molecular mechanisms and selective forces that drive the evolution of NB-LRR (nucleotide binding-leucine-rich repeat) resistance genes. New implications for models of *R* gene evolution have been raised by demonstrations that R proteins can detect cognate Avr proteins indirectly by 'guarding' virulence targets, and by evidence that R protein signalling is regulated by intramolecular interactions between different R functional domains. Comparative genomic surveys of NB-LRR diversity in different species have revealed ancient NB-LRR lineages that are unequally represented among plant taxa, consistent with a Birth and Death Model of evolution. The physical distribution of NB-LRRs in plant genomes indicates that tandem and segmental duplication are important factors in *R* gene proliferation. The majority of *R* genes reside in clusters, and the frequency of recombination between clustered genes can vary strikingly, even within a single cluster. Biotic and abiotic factors have been shown to increase the frequency of recombination in reporter transgene-based assays, suggesting that external stressors can affect genome stability. Fitness penalties have been associated with some *R* genes, and population studies have provided evidence for maintenance of ancient *R* allelic diversity by balancing selection. The available data suggest that different *R* genes can follow strikingly distinct evolutionary trajectories, indicating that it will be difficult to formulate universally applicable models of *R* gene evolution.

* *Correspondence*: Tel.: +1 540 231 2388; Fax: +1 540 231 7126; E-mail: johnmcd@vt.edu

INTRODUCTION

Plants and their pathogens are engaged in a never-ending, convoluted battle of molecular one-upmanship. Pathogenesis schemes are often based upon secretion of large arrays of virulence factors that target molecules inside or outside plant cells and thereby act as molecular saboteurs (Chang *et al.*, 2004; van't Slot and Knogge, 2002). The array of virulence factors deployed by pathogens can be quite variable, even between different strains of the same pathogen species (Greenberg and Vinatzer, 2003). For their part, plants have evolved molecular surveillance systems that provide for recognition of many different types of pathogen-encoded molecules. Plants are capable of recognizing so-called PAMPs (pathogen-associated molecular patterns) that are highly conserved among even distantly related pathogens (reviewed in Nurnberger *et al.*, 2004). However, some pathogens can suppress PAMP-induced defences (e.g. Kim *et al.*, 2005). Thus, plants have also evolved large collections of so-called resistance (R) proteins that recognize specific pathogen virulence factors as signals of invasion [note that a virulence factor, when detected by the host, is genetically redefined as an avirulence (Avr) factor] (Dangl and Jones, 2001; Flor, 1955; Martin *et al.*, 2003). R protein-dependent recognition typically activates a rapid, robust suite of cellular defences, which often include a 'hypersensitive response' (HR) of programmed cell death at the infection site.

This surveillance system is very effective when the plant is fortunate enough to express an R protein(s) that recognizes one or more virulence factors from the prospective pathogen. However, R protein-dependent recognition also represents an Achilles' heel for the plant, because individual R proteins are rendered useless if the pathogen alters or deletes the corresponding avirulence factor. Pathogens are quite adept at this avoidance mechanism, due to multiplicity and redundancy in the virulence factor arsenal. Additional complications for the plant can arise from the capacity of pathogens to 'mask' avirulence factors from detection, or to interfere with defence signalling pathways at multiple points downstream of recognition (Abramovitch and Martin, 2004; Chang *et al.*, 2004). Thus, plants face a considerable challenge in

© 2006 BLACKWELL PUBLISHING LTD

keeping pace with pathogens that deploy versatile, rapidly evolving arrays of weaponry.

R genetic loci often exhibit significant intraspecific polymorphism, suggesting that recognition vs. evasion is a major battleground in the struggle between plants and pathogens (Bergelson *et al.*, 2001; Stahl and Bishop, 2000). For this reason, much effort has been devoted towards understanding how *R* genes evolve. This question is of obvious interest from both applied and fundamental perspectives; indeed, many aspects of current models for *R* gene function and evolution were either foreshadowed or explicitly postulated by plant breeders, based on results from their efforts to deploy *R* genes in crops. In this review, we discuss recent insights into the molecular mechanisms and selective forces that drive *R* gene evolution. This subject is inherently multidisciplinary, and we take three perspectives. First, we adopt a relatively narrow focus, to highlight recent mechanistic insights into R protein functionality that hold implications for *R* gene evolution. We then zoom out for a genome-level perspective, from which we discuss insights provided by comparative genomics. Finally, we take an even broader view that incorporates population-level studies of *R* gene evolutionary dynamics. Our primary emphasis is placed on the most commonly occurring class of *R* genes: the nucleotide-binding, leucine-rich repeat (NB-LRR) superfamily, with reference to other types of *R* genes when appropriate [see Meyers *et al.* (2005) for an up-to-date summary of R protein structure]. We emphasize advances from the last 5 years, and attempt to predict areas of future interest.

A MOLECULAR PERSPECTIVE: GUARDS AND TRIGGER GUARDS IN THE PLANT SURVEILLANCE SYSTEM

In order to have a complete understanding of how *R* genes evolve, it is necessary to understand what R proteins actually recognize and how this recognition occurs at a molecular level. H. H. Flor's gene-for-gene hypothesis predicts that resistance occurs when an *R* gene recognizes a specific *Avr* gene in the pathogen (Flor, 1955). A long-standing molecular interpretation of this genetic model is that *R* genes encode receptors that directly bind the cognate Avr proteins. This model is based on analogy to the receptor–ligand relationship of antibodies and antigens, and is consistent with genetic specificity of R–Avr interactions and the allelic diversity that often occurs at *R* and *Avr* loci. A putative molecular platform for R–Avr interactions was identified during the initial molecular cloning of *R* genes, which revealed that LRRs are a signature motif of R proteins (Jones and Jones, 1996). LRRs in other proteins had previously been shown to form a beta sheet that provides a versatile and highly evolvable ligand-binding surface (Kobe and Deisenhofer, 1994; Kobe and Kajava, 2001). LRR-encoding regions in many NB-LRR genes are subject to strong diversifying selection, as evidenced by high ratios of non-synonymous to synonymous nucleotide substitutions (e.g. Botella *et al.*, 1998; McDowell *et al.*, 1998; Meyers *et al.*, 1998; Mondragon-Palomino *et al.*, 2002; Parniske *et al.*, 1997). Moreover, domain-swapping experiments provided direct evidence that LRRs play a major role in recognition specificity (e.g. Ellis *et al.*, 1999; Thomas *et al.*, 1997). These data were generally interpreted as evidence that LRRs are under selection to bind directly different avirulence factors (again inspired by analogy to self–non-self recognition systems in animals (e.g. Kreitman and Akashi, 1995)).

Although this hypothesis provides a simple, appealing parallel to our own immune system, it is not strongly supported by molecular assays for direct interaction between NB-LRR proteins and cognate Avr proteins. Direct R–Avr interactions have been demonstrated in only three studies (Deslandes *et al.*, 2003; Dodds *et al.*, 2006; Jia *et al.*, 2000), whereas the majority of such tests have been negative (e.g. Luderer *et al.*, 2001). Moreover, Avr-binding activities have been demonstrated in plants that are null for the corresponding R protein, suggesting the existence of intermediaries in R–Avr interactions (e.g. Kooman-Gersman *et al.*, 1996). Finally, as mentioned above, it is increasingly clear that Avr determinants actually promote virulence in susceptible plants. These observations were synthesized in a ground-breaking commentary which proposed that R proteins 'guard' plant proteins that are targeted by cognate Avr proteins to promote virulence (Van der Biezen and Jones, 1998). If the guard (R protein) senses that its guardee (virulence target) has been attacked, then cellular alarms are sounded and defences are activated. However, if the guard is absent (not expressed) or incompetent (unable to perceive modification of the virulence target), then colonization proceeds without interference. In other words, the R protein detects the virulence-promoting activity of the cognate Avr, rather than the Avr itself (see excellent reviews in Innes, 2004; Van der Hoorn *et al.*, 2002).

The *Arabidopsis R* genes *RPM1* and *RPS2* serve as exemplars for this model. Both of these *R* genes respond to *Pseudomonas syringae* virulence factors, termed 'effectors', that are exported to the interior of plant cells by Type III secretion (Gopalan *et al.*, 1996; Leister *et al.*, 1996; Nimchuk *et al.*, 2000). These effectors do not bind to RPM1 or RPS2, rather they target a novel, membrane-associated protein called RIN4 that physically co-localizes with RPM1 and RPS2 (Axtell and Staskawicz, 2003; Kim *et al.*, 2005; Mackey *et al.*, 2002, 2003). RPM1 and RPS2 are activated in response to effector-dependent modification or degradation of RIN4. Equally compelling examples of guarding are provided by the *Arabidopsis* RPS5 protein, which detects cleavage of an *Arabidopsis* Ser–Thr kinase by an Avr-encoded protease (Shao *et al.*, 2003), and by the tomato Cf-2 protein, which detects inhibition of a tomato protease by a pathogen-encoded protease inhibitor (Kruger *et al.*, 2002; Rooney *et al.*, 2005).

Collectively, these examples indicate that a significant proportion of R proteins conform to the guard model. However, it is

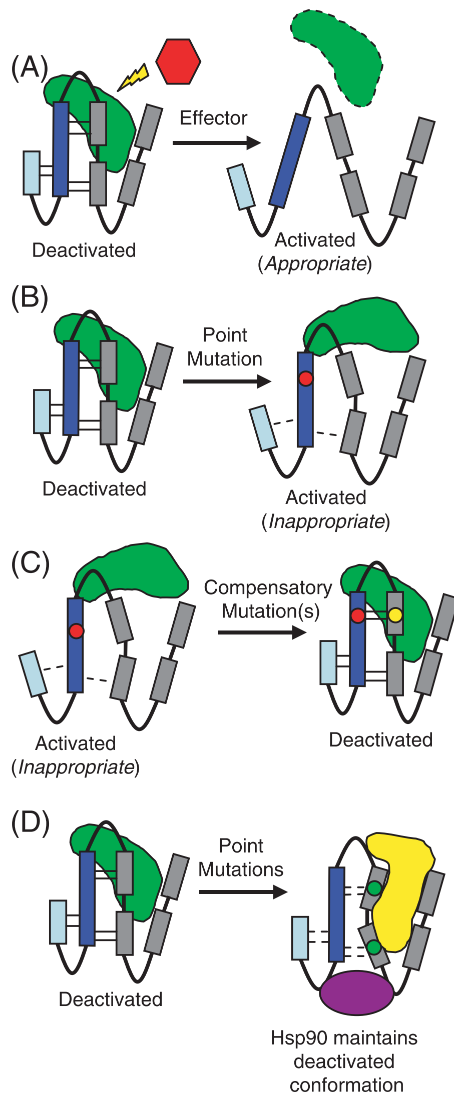

Fig. 1 Speculative models depicting evolutionary implications of repressive intramolecular interactions ('RInIs'), and HSP90 association. (A) R proteins are negatively regulated in part by RInIs (horizontal lines) between the N-terminal domain (light blue), nucleotide binding site (dark blue) and leucine-rich repeats (grey). Intermolecular interaction with the virulence target/guardee (green) may also contribute to repression. Effector-dependant modification or degradation of the guardee disrupts RInIs, thereby appropriately activating the R protein and initiating defence responses (the schematic is modified from Belkhadir *et al*., 2004). (B) RInIs are disrupted by a spontaneous point mutation in the NBS (red circle), leading to inappropriate defence activation in the absence of pathogen attack. Note that activation need not be 'full' to have deleterious effects. (C) RInIs are restored to the mutant from B due to a compensatory mutation(s) in an LRR (yellow circle), restoring the ability to assume a deactivated conformation. (D) Mutations in the LRRs (green circles) enhance a new binding specificity (ligand shown in yellow) but destabilize RInIs. This mutant is stabilized in an inactive, non-deleterious conformation by interaction with HSP90 (purple). Subsequent compensatory mutations in the LRRs or other domains (as shown in C) could restore RInIs and enhance the stability of repression.

important to bear in mind that direct interactions between two NB-LRR proteins and cognate Avr proteins have been documented, indicating that some R proteins conform to the classic receptor–elicitor model (Deslandes *et al*., 2003; Dodds *et al*., 2006; Jia *et al*., 2000). Further experimentation will be necessary to determine the relative prevalence of guarding vs. direct detection. As noted previously, the NB-LRR genes that have been shown to act as guards appear to be evolving conservatively (McDowell, 2004). It will be interesting to determine the prevalence of guarding among NB-LRR genes that are under strong diversifying selection.

Although the guard model may not explain all R–Avr interactions, it undoubtedly explains a significant proportion and therefore holds important implications for models of R–Avr coevolution. For example, it is reasonable to predict that different pathogens might target the same guardees (e.g. regulators of defence). Thus, plants could evolve multispecificity R proteins that guard popular virulence targets. Guarding thereby provides the plant with a mechanism to recognize a large number of pathogen effectors using a relatively small number of *R* genes. For example, *RPM1* recognizes two sequence-unrelated effectors through their effects on *RIN4* (Mackey *et al*., 2002). It is also plausible that different plant species could independently evolve *R* genes that guard the same virulence target. One such case of potentially convergent evolution was recently reported by Ashfield and colleagues, in which a functionally equivalent but independently evolved analogue of *RPM1* was identified in soybean. Preliminary evidence suggests that this gene might guard the soybean *RIN4* orthologue (Ashfield *et al*., 2004).

Another major implication of the guard model relates to what R proteins are under evolutionary pressure to recognize. Rather than evolving to bind new or different effectors, nascent R proteins might be under selective pressure to bind or otherwise monitor newly 'at-risk' virulence targets. If true, then selection might also favour mutations in the virulence targets themselves to optimize interaction with a prospective guard. Moreover, it might be to the plant's advantage to evolve 'decoys', proteins that structurally mimic a virulence target but do not play a role in normal plant processes. Such decoys might substitute for virulence targets in R protein complexes, thereby allowing the virulence target to carry out its normal function in plant cells, free from structural or functional constraints necessitated by interaction with a guard.

On this note, it is relevant to consider another recent molecular insight with evolutionary implications: R protein activity appears to be negatively regulated by repressive intramolecular interactions between the NB, LRR and N-terminal domains (termed 'RInIs', and shown diagrammatically in Fig. 1A; reviewed in Belkhadir *et al*., 2004; Innes, 2004; Rathjen and Moffett, 2003). This concept is supported by two lines of evidence. First, several recent studies have reported gain-of-function mutations in *R*

genes that confer constitutive resistance activation in the absence of pathogens (Fig. 1B; reviewed in Belkhadir *et al.*, 2004). These mutants typically display a stunted growth habit and spontaneous cell death. Second, a highly original study by Moffett *et al.* (2002) demonstrated that different domains of the Rx and Bs2 proteins, when expressed in *trans* from different transgenes, are capable of associating with each other *in vivo* and reconstituting a functional resistance specificity. Importantly, interactions between the different domains are disrupted by the cognate Avr. Thus, intramolecular interactions are thought to maintain R proteins in a signalling-competent but inactive conformation. This conformation is altered directly or indirectly by the corresponding Avr, leading to engagement of downstream signalling components and activation of defences (Fig. 1A). The virulence target can be incorporated into this model as an inhibitory partner in a 'trigger complex', whereby Avr-dependent alteration or loss of the virulence target could disrupt intramolecular interactions and activate defence signalling (Belkhadir *et al.*, 2004). Put another way, inhibitory intra- and/or intermolecular interactions could serve as 'trigger guards', analogous to the metal loop that surrounds the trigger of a firearm to reduce accidental discharge.

This model should be viewed with caution because important aspects have not been experimentally confirmed (e.g. structural confirmation that gain-of-function mutants are activated due to disruption of inhibitory interactions). However, if true, this model holds important implications for *R* gene evolution. The first is that R proteins are inherently dangerous to the plant, due to their capacity for spontaneous activation of defences in the absence of pathogen invasion. This is not a new idea (see below), but recent demonstrations that R proteins can be improperly activated by simple point mutations strikingly underlines the risk associated with expressing dozens or hundreds of NB-LRRs on a continuous basis. The second implication is that LRRs may actually play a dual role as recognition specificity determinants and as trigger guards. As mentioned above, LRRs are under diversifying selection, presumably to bind new specificity determinants. However, recognition specificity determinants have also been mapped outside the LRRs (Ellis *et al.*, 1999; Luck *et al.*, 2000). Moreover, certain codons outside the LRRs are under diversifying selection (e.g. Luck *et al.*, 2000; Mondragon-Palomino *et al.*, 2002). Taken together, these observations suggest that LRRs may be coevolving with other domains in the R protein to optimize recognition specificity while maintaining necessary inhibitory interactions. Evolution of new interaction specificity might sometimes involve substitutions in the LRRs or other domains that enhance recognition but disrupt intramolecular interactions and permit constitutive activity. Selection may then drive compensatory mutations in the LRRs and/or in other domains to restore trigger guard function while maintaining the new recognition capability (Fig. 1C,D). This hypothesis can be addressed as we achieve higher resolutions of R protein structure.

In this context, it is relevant to consider recent reports of association between NB-LRR proteins and the molecular chaperone HSP90. HSP90 proteins associate with a select but diverse array of eukaryotic client proteins, many of which function as receptors or downstream components of signal transduction pathways (reviewed in Mayer and Bukau, 1999). Several plant R proteins have now been associated with HSP90 (reviewed in Schulze-Lefert, 2004). By analogy to the role of HSP90s in mammalian steroid hormone receptor complexes, plant HSP90s may stabilize R proteins in a signalling-competent conformation, but the exact role(s) of HSP90 in R protein functionality remain to be defined. From an evolutionary perspective, the addition of R proteins to the HSP90 clientele is intriguing in light of recent evidence that HSP90s act as buffers that modulate the phenotypic consequences of genetic variability. This newly discovered role is supported by studies in which previously unseen phenotypic variability was revealed when HSP90 was impaired genetically, pharmacologically, or by environmental stress (reviewed in Sangster *et al.*, 2004). According to current models, HSP90 can mask the effect of mutations by several potential mechanisms, for example by stabilizing mutant proteins in a conformation approximating that of the wild-type protein. Thus, mutations that might otherwise be deleterious are not subjected to purifying selection, providing a potential window of opportunity for additional mutations to accrue. Accordingly, HSP90s could facilitate *R* gene evolution by stabilizing variants that provide new recognition, but are inherently unstable or deleterious due to disruption of intramolecular interactions (Fig. 1D). In other words, HSP90-dependent stabilization could thereby lengthen the window of opportunity for potentially useful R protein variants to evolve an optimal conformation (Sangster *et al.*, 2004; Schulze-Lefert, 2004).

A GENOMIC PERSPECTIVE: PROLIFERATE, RECOMBINE (OR NOT), AND CONQUER

A wealth of information about NB-LRR evolution has been provided by whole genome sequences, ESTs, and PCR-based surveys of NB-LRR diversity. At the most fundamental level, genome surveys have confirmed that all plants maintain large, diverse NB-LRR families. For example, the landmark paper by Meyers *et al.* (2003) provided the first complete characterization of a plant NB-LRR complement, which consists of 149 NB-LRR genes in the Columbia ecotype of Arabidopsis. Other plants appear to maintain even larger NB-LRR families (e.g. Monosi *et al.*, 2004; Pan *et al.*, 2000). To date, no function other than disease resistance has been described for any plant NB-LRR gene. Thus, a significant proportion of the plant genome appears dedicated to NB-LRR-based pathogen surveillance, and NB-LRR gene duplication has had a prodigious impact on the evolution of plant immunity.

Sequence comparisons have also provided clear evidence for ancient lineages within the NB-LRR superfamily. For example, two NB-LRR subfamilies have been defined based on domains at the amino-terminal end of the encoded protein, which contain either a signature domain from the Toll/interleukin-1 family of receptors that mediates innate immunity in animals, or a coiled-coil region (referred to, respectively, as TNL and CNL, following Meyers *et al*., 1999). TNLs and CNLs are found in dicots and gymnosperms, suggesting that these subfamilies are ancient and probably evolved before the angiosperm/gymnosperm split. A large-scale phylogenetic analysis by Cannon *et al*. (2002) further partitioned the CNL subfamily into at least four distinct lineages, some of which also pre-dated the angiosperm/gymnosperm divergence.

Lineages within the NB-LRR superfamily are not equally represented among all plant taxa. The most extreme example is that TNL genes are absent from every grass genome surveyed to date (e.g. Bai *et al*., 2002). TNLs are present in pine and moss, indicating that the TNL subfamily evolved prior to the angiosperm–gymnosperm split and was lost during monocot evolution (Akita and Valkonen, 2002; Meyers *et al*., 1999, 2002). By contrast, TNL genes are generally overrepresented, compared with CNLs, in dicot genomes. For example, there are nearly twice as many TNL genes as CNL genes in the Arabidopsis genome (Meyers *et al*., 2003). The conserved intron positions of the TNLs and the shorter phylogenetic branch lengths within the TNL phylogeny suggest a shorter evolutionary history (Cannon *et al*., 2002; Meyers *et al*., 2003). The ancient CNL lineages are also unequally represented in dicot taxa. For example, one CNL lineage is represented by 42 sequences from *Medicago truncatula* and soybean, but only two from Arabidopsis, while another lineage is represented by 18 Arabidopsis genes but only four from *M. truncatula* and soybean (Cannon *et al*., 2002). This unequal representation is consistent with a 'birth and death' model in which some NB-LRR lineages have been retained for relatively long time spans, whereas others have been lost and supplanted by more recently emergent lineages (Michelmore and Meyers, 1998; Nei and Rooney, 2005). The uneven distribution of NB-LRR lineages within plant species also underlines that no single plant genome provides a universally applicable model of NB-LRR diversity (Cannon *et al*., 2002).

The large size of NB-LRR families raises another question: What are the mechanisms through which NB-LRR genes proliferate? As predicted from classical genetic characterization of *R* loci, NB-LRR genes are often organized as clusters. For example, 109 of the 149 Arabidopsis NB-LRR genes reside in 40 clusters ranging in size from two to eight genes, while the remaining 40 genes exist as singletons (Meyers *et al*., 2003; Richly *et al*., 2002). The rice genome contains an approximately equivalent proportion of clustered genes (Bai *et al*., 2002; Monosi *et al*., 2004; Zhou *et al*., 2004). This physical clustering suggests that tandem duplication is an important source of new *R* genes, and this hypothesis is supported by two trends. First, genes within a cluster tend to occupy the same phylogenetic lineage (but see below). Second, gene copy number can vary widely among haplotypes within a species. One extreme example is provided by haplotypes of the maize *Rp1* cluster, in which a range of 1–52 copies has been estimated (Smith *et al*., 2004). Such variation in gene copy number can be driven by mispairing and unequal crossing over between linked NB-LRR genes (intragenic recombination) or between repetitive sequences such as transposons in intergenic regions (intergenic recombination) [see Leister (2004) for an excellent explanation of recombination-associated terminology]. Evidence for both has been provided (reviewed in Ellis *et al*., 2000; Hulbert *et al*., 2001).

Although local, tandem duplication explains the origin of a large fraction of NB-LRRs, genome surveys have revealed 'mixed clusters' comprising evolutionarily distant NB-LRR genes. For example, at least ten clusters in Arabidopsis contain interspersed CNLs and TNLs (Meyers *et al*., 2003; Richly *et al*., 2002). Such distributions obviously could not have arisen through tandem duplications beginning from a single progenitor gene. Conversely, some closely related *R* genes are physically dispersed in the genome (i.e. on different chromosomes). Two models have been proposed to explain NB-LRR distribution patterns that are inconsistent with simple tandem duplication (reviewed in Leister, 2004). The 'rapid rearrangement' model emphasizes the importance of ectopic duplications, in which single genes or small groups are transposed to distal locations by an undefined mechanism. This model is supported by an apparent lack of large-scale synteny in sequences flanking related NB-LRR genes in cereals and Arabidopsis (Leister *et al*., 1998; Meyers *et al*., 2003; Richly *et al*., 2002). A different but non-exclusive, 'conserved synteny' model emphasizes the importance of large-scale segmental duplication with subsequent local rearrangement. These models were evaluated in an intriguing study which utilized statistical approaches from phylogenetics and phylogeography in which chromosomal regions were considered as geographical populations, from which migration patterns could be evaluated (Baumgarten *et al*., 2003). This analysis indicated that the extant complement of 149 NB-LRRs in Arabidopsis could be explained by 89 duplications, 71 of which were 'local' (i.e. the duplication products resided within 2 Mb of each other). Of the 18 non-local duplications, 15 could be explained by large-scale segmental duplications and only three may have involved ectopic duplication. This analysis supports the conserved synteny model. However, the authors' definition of 'cluster' (2-Mb regions) differed substantially from the definition used in previous analyses (two of eight contiguous genes) that supported the prevalence of rapid rearrangement (Meyers *et al*., 2003; Richly *et al*., 2002). As pointed out by Leister (2004), some of the 2-Mb clusters contain gene duplications that are relatively distant from each other and

are non-syntenic in flanking regions. The mechanisms that generate these local, non-tandem duplications remain to be determined, and are likely to be an important factor in *R* gene proliferation. Another question of interest relates to why and how some R genes proliferate in clusters, whereas related genes are restricted as single-copy loci.

Given the prevalence of *R* gene clusters, to what extent is *R* gene structure and function influenced by sequence exchanges between duplicated genes through unequal crossing over or gene conversion? A large amount of genetic and molecular evidence indicates that sequence exchanges are a significant factor in *R* gene cluster evolution (reviewed in Ellis *et al*., 2000; Hulbert *et al*., 2001). However, some *R* gene lineages appear to be evolving in relative isolation with little or no sequence exchanges, raising questions about the relevance of sequence exchanges in *R* gene evolution (Ellis *et al*., 2000; Michelmore and Meyers, 1998). This issue has been addressed by recent demonstrations that the relative importance of sequence exchange varies not only between different *R* loci, but between genes within individual clusters. Moreover, different genes within a cluster can evolve at very different rates. The first such example of heterogeneous evolution was provided by the *N* cluster for flax rust resistance (Dodds *et al*., 2001). Genes within the *N* cluster fall into three distinct subclasses that do not recombine with each other. However, sequence exchanges appear frequent within one subclass containing two paralogues that share a high degree of sequence identity. More recently, Kuang *et al*. (2004) undertook an extremely ambitious comparison of haplotype diversity at the *RGC2* cluster in cultivated lettuce and wild relatives. The *RGC2* cluster is large, complex and variable in terms of copy number (12–32). Kuang and colleagues cloned and analysed hundreds of *RGC2* genes from dozens of *Lactuca* accessions. This unprecedented depth of sampling allowed clear delineation of two types of *RGC2* genes: so-called Type I genes, which are evolving rapidly and display chimeric structures indicative of frequent sequence exchange, and Type II genes, which are relatively conserved and appear to recombine very infrequently. A subsequent analysis of the potato *R1* NB-LRR cluster revealed three subclusters of Type I genes, with frequent exchanges within but not between subclusters (Kuang *et al*., 2005).

Why is it that related, physically contiguous genes have such different evolutionary modes and tempos? Kuang *et al*. (2004) speculated that Type I genes can be stochastically converted into Type II genes by structural rearrangements that inhibit mispairing and thereby reduce sequence exchange. It is also tempting to speculate that at least a subset of Type II genes are conserved because they have evolved important resistance specificities (Kuang *et al*., 2005). By contrast, some Type I genes might be works-in-progress. Consistent with this idea, *RGC2* Type I genes appear to be subject to strong diversifying selection while Type II genes are under purifying selection (Kuang *et al*., 2004).

The studies above also speak to the question of how useful genes and haplotypes might be protected from homogenization that could result from repeated sequence exchanges between related genes (concerted evolution) (Nei and Rooney, 2005). It is to be expected that duplicated genes are less likely to recombine as they independently accumulate nucleotide substitutions (Dodds *et al*., 2001). This is supported by analyses of NB-LRRs in *Arabidopsis*, in which exchanges between highly similar genes appear relatively frequent, whereas exchanges between more divergent sequences are rare (Baumgarten *et al*., 2003; Meyers *et al*., 2003; Mondragon-Palomino *et al*., 2002). In some cases intergenic regions have diverged considerably between haplotypes, perhaps to inhibit mispairing and thereby stabilize advantageous haplotypes (e.g. Parniske *et al*., 1997). As pointed out by Baumgarten, dispersal of duplicated copies to physically distant sites (regardless of whether this occurs by segmental duplication or ectopic recombination) provides another mechanism by which useful genetic novelty could escape erasure by concerted evolution (Baumgarten *et al*., 2003; but see Parniske and Jones, 1999, for an interesting exception). Finally, Kuang *et al*. made the very interesting observation that introns in LRR-encoding regions of *RGC2* and *R1* Type I genes are more similar to each other than are the flanking exons. This pattern indicates that diversifying selection counteracts concerted evolution. This emphasizes the strength of selection for evolutionary novelty at Type I *R* loci.

It has long been speculated that the frequency of recombination could be elevated in response to unfavourable environmental conditions, thereby increasing the probability that advantageous genetic novelties will arise (McClintock, 1984). This hypothesis has been supported by recent studies that use genetic screens to provide quantitative measures of recombination frequency in response to pathogen colonization and other stresses. These screens are based on assays for restoration of a split *LUC* or *GUS* reporter transgene by unequal crossing over. Such crossovers can be visualized as sectors of reporter gene activity in leaves that can be counted as a measure of somatic recombination frequency. Lucht *et al*. (2002) used this approach to demonstrate a 1.8-fold increase in somatic recombination after Arabidopsis plants were infected with *Hyaloperonospora parasitica*. Additionally, they demonstrated elevated recombination in plants treated with exogenous salicylic acid and in a mutant with constitutively active defences. Kovalchuk *et al*. (2003) observed elevated recombination of a reporter transgene and an endogenous gene in non-inoculated leaves of tobacco plants infected by tobacco mosaic virus. This suggested the presence of a 'systemic recombination signal' (SRS) that can spread to non-infected tissues of the plant independently of the pathogen. Further evidence for an SRS was shown with grafting experiments demonstrating an increase in recombination in plants that were grafted with an inoculated, 'signal-carrying' leaf as compared with plants grafted with a non-inoculated leaf. A subsequent

study addressed whether an increase in recombination is a pathogen-specific response or whether it can be triggered by multiple stressors. Filkowski *et al.* (2004) demonstrated that the radical-generating agents UVC or Rose Bengal significantly increased local and systemic recombination, suggesting that recombination rates could be elevated by any stress associated with production of reactive oxygen intermediates. The molecular identity of the SRS is currently unknown.

What might be the adaptive advantage for the plant to increase the frequency of somatic recombination? Given the probable importance of recombination in *R* gene evolution, it is certainly tempting to speculate that elevated somatic recombination could accelerate the evolution of new *R* genes. Although somatic recombination might not generate useful variation in short-lived species, it could be important for long-lived plants or species that reproduce clonally. Moreover, somatic recombination events can be transmitted to progeny if they occur in tissue that gives rise to reproductive structures. Accordingly, Kovalchuk *et al.* (2003) reported a three-fold increase in LUC+ progeny in infected plants, indicating increased frequency of meiotic and/or heritable somatic recombination. It will be of great interest to determine whether meiotic recombination frequency is indeed elevated in response to pathogens, to assay effects of induced recombination on the evolution of clustered *R* genes and to dissect the mechanisms that promote stress-induced recombination.

These studies raise the issue of whether recombination can generate new recognition specificity over short time frames. Naturally occurring, interallelic recombination at the flax rust resistance locus *L* can create chimeric genes encoding novel proteins with race-specificities that differ from both parental alleles (Ellis *et al.*, 1999; Luck *et al.*, 2000). While underlining the capacity of recombination to generate useful variants quickly, these studies highlight our current lack of knowledge about the relative frequency and functional importance of genetic exchange between alleles vs. paralogues (Michelmore and Meyers, 1998). Analysis of the maize *Rp1* locus has revealed another potential mechanism whereby intergenic recombination can quickly generate new recognition. Screens of the *Rp1* cluster have identified mutants that exhibit non-parental race specificities (Hulbert *et al.*, 2001). Based on the precedent from *L* discussed above, it would be expected that most of these variants arose from creation of new chimeric genes. Surprisingly, recent molecular analysis of four *Rp1* variants did not reveal any chimeric genes (Smith and Hulbert, 2005). Rather, it appears that intergenic recombination has created reassorted haplotypes with a new specificity driven by two or more genes/alleles that were not associated in the parents of the cross. The genes responsible for this new specificity have not been identified, but this study brings to mind tantalizing analogies to metazoan Toll-like receptors, in which recognition specificity is expanded by heterodimeric interactions (Kawai and Akira, 2005).

A POPULATION PERSPECTIVE: BALANCING THE COSTS AND BENEFITS OF SURVEILLANCE

Notable insights into *R* gene evolutionary dynamics have been provided by recent studies that integrate concepts and resources from molecular genetic and genomic studies into population genetic and ecological contexts. To begin with, plant–pathogen coevolution is often described as 'dynamic', 'rapidly evolving', etc., implying rapid turnover of alleles at both *R* and *Avr* loci. However, recent population-level studies have revealed the existence of remarkably long-lived polymorphism at *R* gene loci. The first example was provided by surveys of allelic diversity at the Arabidopsis *RPM1* locus (Stahl *et al.*, 1999). In *A. thaliana*, *RPM1* is a single-copy gene in resistant ecotypes, and is completely absent from susceptible (*rpm1*) ecotypes. The *RPM1* loci in *Brassica oleracea* and *Arabidopsis lyrata* are syntenic with functional *RPM1* in Arabidopsis, demonstrating that the *A. thaliana rpm1* null allele was created by deletion of a functional *RPM1* gene (Grant *et al.*, 1998; Stahl *et al.*, 1999). Analysis of nucleotide divergence in the regions flanking the *A. thaliana rpm1* null allele indicates that this allele has been maintained for approximately 10 million years, placing its origin close to the time of Arabidopsis speciation (Stahl *et al.*, 1999). Notably, this ancient origin is inconsistent with a classic 'arms race' model, which predicts that defeated *r* alleles are supplanted by new *R* alleles through selective sweeps, in which the obsolete alleles are removed from the population. Rather, modelling studies by Stahl *et al.* (1999) support a 'trench warfare' model for *RPM1* evolution, in which the functional and null alleles are long-lived, but their relative frequencies fluctuate dynamically over time, due to recurrent cycles of negative frequency-dependent selection.

Why then might a null allele of a resistance gene with obvious selective advantages (recognition of two effectors) have been maintained for such a long time? One explanation is that the functional *RPM1* allele is selectively disadvantageous in the absence of pathogen pressure. Recent studies under laboratory and field conditions have demonstrated fitness penalties associated with general inducers of resistance or gain-of-function mutants in signalling components that confer constitutive resistance (e.g. Heidel *et al.*, 2004; reviewed in Heil, 2002). These studies underline that resistance, when induced, can be costly, but do not address the costs associated with maintenance of large numbers of *R* genes. As discussed above, these molecular sentries have hair triggers, and the trigger guards appear to be far from foolproof (Fig. 1). However, fitness costs associated with individual *R* genes have been difficult to quantify because of a lack of truly isogenic resistant and susceptible lines that differ only in the genotype at a single *R* locus. Tian *et al.* (2003) overcame this problem with a clever strategy in which an *rpm1* null line of Arabidopsis was transformed with a functional *RPM1* transgene flanked by LOX recognition sites for the CRE recombinase. The

LOX sites enabled excision of the *RPM1* transgene through crosses to plants expressing a CRE transgene, thereby creating matched sets of transgenic lines that differ only in the presence or absence of *RPM1*. Analysis of four independently derived transgenic sets, under field conditions with no obvious pathogen pressure, revealed that the *RPM1* transgene conferred a reproducible and remarkably high 9% decrease in seed production. This result is entirely consistent with a model in which balancing selection has driven long-term maintenance of a null *rpm1* allele to counterbalance a potent but costly *RPM1* allele.

How generally applicable is this example? Interestingly, the Arabidopsis *RPS5* locus also seems to have been maintained as a presence/absence polymorphism (Henk *et al*., 1999). Moreover, Nobuta *et al*. (2005) have provided evidence that duplicated *R* genes have been frequently deleted during the evolution of the Arabidopsis genome, suggesting that superfluous *R* genes might carry a cost. However, as noted by Tian and others, it seems unlikely that many *R* alleles are as costly as *RPM1*, because their cumulative load would be unsustainable (Brown, 2003; Tian *et al*., 2003). Indeed, a subsequent study that assessed fitness costs associated with the Arabidopsis *R* genes *RPS2* (for resistance to *Pseudomonas syringae*) or *RPP5* (resistance to the downy mildew pathogen *Hyaloperonospora parasitica*) revealed a complex and counterintuitive scenario (Korves and Bergelson, 2004). Neither of these *R* genes was costly in the absence of disease pressure. In fact, susceptible lines displayed slightly elevated fitness, relative to resistant lines, under conditions of disease pressure. The authors termed this phenomenon 'net cost of resistance under attack'. In the case of *RPP5*, a fitness cost was apparent only when competition (high plant density) was built into the experiment, suggesting a link to nutrient availability. Interestingly, the fitness of susceptible lines appeared to be elevated by a developmental compensation whereby diseased, susceptible plants flowered more quickly (Korves and Bergelson, 2003, 2004). These studies emphasize that cost–benefit relationships for individual *R* genes may vary substantially, according to the nature of allelic variability at the *R* locus in question, developmental responses to the pathogen, availability of nutrients and environmental conditions. It will be necessary to test a large number of additional *R* alleles in diverse species, under relevant environmental conditions, to determine whether generalizations about *R* gene-dependent fitness costs are possible. It will also be of interest to determine the molecular basis of the *RPM1* fitness cost. Is this gene draining plant resources by triggering weak defence activation in the absence of pathogens? Might there be additional, naturally variable loci that buffer the negative effects of *RPM1*?

Studies that followed Stahl's examination of *RPM1* provided additional evidence for long-lived polymorphisms at a variety of *R* gene loci, indicating that balancing selection is of general importance for *R* gene evolution (Allen *et al*., 2004; Bergelson *et al*., 2001; Kuang *et al*., 2004; Mauricio *et al*., 2003; Tian *et al*., 2003). At present, it is not clear whether a cost of resistance is relevant in any of these cases. It is important to recognize that it is unnecessary to invoke a cost of resistance to explain all long-lived *R* polymorphisms. Balancing selection may also act to maintain alternative alleles with different recognition specificities. The benefit of such cognitive diversity is intuitively obvious and is supported by modelling studies (reviewed in Brown, 2003). Interestingly, the Arabidopsis *RPP13* gene and the corresponding downy mildew *Avr* gene are extremely polymorphic and are under strong diversifying selection (Allen *et al*., 2004; Rose *et al*., 2004). A similar pattern is evident between alleles at the flax *L* locus and corresponding *Avr* alleles in flax rust (Dodds *et al*., 2006). These studies are the first examples in which both the host *R* alleles and the corresponding *Avr* alleles are subject to strong diversifying selection, highly suggestive of reciprocal coevolution. It is tempting to speculate that these R proteins interact directly with the corresponding Avr, and that this coevolution is driven by selection of new *Avr* alleles to avoid this interaction and subsequent selection of new *R* alleles that have re-established contact. Interestingly, the allelic diversity at *RPP13* appears relatively ancient, while the diversity at *L* emerged relatively recently (Bergelson *et al*., 2001).

CONCLUSION: CELEBRATE DIVERSITY

We hope to have adequately conveyed the exciting advances in understanding of *R* gene evolution that have occurred over the last 5 years, and to have highlighted important unresolved issues. We are particularly intrigued by the apparent diversity in evolutionary trajectories of different NB-LRR genes. For example, *RPM1* and *RPS5* appear relatively stable: functional alleles of *RPM1* and *RPS5* are subject to purifying selection, and allelic diversity at both loci is very limited and of relatively ancient origin. Both of these R proteins recognize important pathogen effectors indirectly, by guarding putative virulence targets. Thus, these loci seem to exemplify cases in which arms races (if any) were resolved long ago and were followed by recurrent cycles of frequency-dependent selection. By contrast, *RPP13* and *L* are influenced by frequent interallelic recombination and strong diversifying selection, which has given rise to extensive allelic polymorphism. Moreover, the corresponding *Avr* loci are also under diversifying selection, which is highly suggestive of reciprocal coevolution driven by selection for (and against) direct interaction between the R and Avr proteins. Evidence for direct, allele-specific interaction between L variants and cognate Avr proteins has recently been obtained (Dodds *et al*., 2006). Thus, *RPP13* and *L* differ substantially from exemplars of the guard model (i.e. RPM1 and RPS5) in their mode and tempo of evolution, and perhaps in their mode of recognition. Together, these examples underline the difficulty of formulating universally

applicable models for *R* gene evolution. It is tempting to speculate that the ubiquity of NB-LRR proteins as pathogen surveillance components can be explained by their collective versatility in recognizing pathogens 'directly' via Avr binding, or indirectly via guarding.

We find it particularly exciting that advances in understanding of *R* gene evolution have been driven to some extent by interdisciplinary cross-fertilization, and we expect more of this in the future. For example, resolution of the structural details of R protein complexes will pave the way for a much more sophisticated evolutionary interpretation of nucleotide substitution patterns in *R* gene sequences. Conversely, deeper sampling of allelic diversity in natural populations for evolutionary studies may enable a higher resolution of residues important for functionality (e.g. using statistical approaches to identify residues in different R protein domains that could be coevolving to maintain optimal tertiary/quaternary structure, Fig. 1). Continued interdisciplinary inquiry will, we hope, reveal new principles useful for breeding durable resistance by conventional or transgenic means, thereby closing the circle of inquiry into *R* gene evolution that was initiated decades ago by plant breeders.

GLOSSARY

Balancing selection—Natural selection that maintains allelic polymorphism within a population.

Convergent evolution—Refers to instances in which the same trait arises *independently* in different evolutionary lineages.

Diversifying selection—Natural selection that favours genetic variation. An elevated rate of non-synonymous substitutions (resulting in amino acid changes), compared with synonymous substitutions, can be interpreted as evidence for diversifying selection.

Fitness—Ability to survive and reproduce. Fitness penalties reduce reproductive capacity (e.g. seed production).

Gene conversion and crossing-over—Gene conversion events result in non-reciprocal genetic exchange, in which the sequence from one strand of DNA is transferred to another DNA strand. By contrast, cross-over events result in the reciprocal exchange of DNA sequences between strands.

Negative frequency-dependent selection—Natural selection that favours rare phenotypes. For example, a new (rare) *R* allele might confer an initial selective advantage. As this allele becomes more prevalent in the plant population, natural selection on the pathogen will favour the emergence of strains that have altered or discarded the Avr gene. As these variants emerge and become prevalent, the 'defeated' *R* allele might become disadvantageous, particularly if the *R* allele reduces fitness in the absence of pathogens.

Paralogue and ortholog—The term paralogue designates different genes that arose through a duplication (e.g. by unequal crossing over). The term orthologue refers to the same gene in different species.

Phylogeography—A subdiscipline of evolutionary biology that uses molecular phylogenetic approaches to understand factors that influence geographical dispersal of populations within a species.

Polymorphism—Naturally occurring variation in sequence or function between alleles of a gene. For example, different alleles of an *R* gene may provide resistance to different pathogen isolates.

Purifying selection—Natural selection against genetic variation.

Segmental duplication—Genetic duplication of a chromosomal segment that includes multiple genes. This can give rise to 'blocks' of duplicated genes that reside at unlinked genetic loci.

Somatic recombination–Genetic recombination that occurs during mitosis in non-reproductive tissue. The products of somatic recombination can be transmitted to progeny only if the recombination occurs in a cell lineage that subsequently develops into reproductive tissue.

Synteny–Conservation of gene order between species.

Tandem and ectopic duplication–Tandem gene duplications give rise to physically proximal duplicated genes. Ectopic duplications result in gene copies that are physically distal (e.g. genetically unlinked).

Unequal crossover–Reciprocal recombination that can result from mispairing and crossing over between genes that are different but share sequence similarity. For example, unequal crossovers can occur between members of a tandemly arrayed multigene family, giving rise to a gene duplication on one chromosome (i.e. a tandem duplication) and a deletion on the other.

ACKNOWLEDGEMENTS

We thank Dr Blake Meyers and anonymous reviewers for insightful comments. Research in the McDowell lab is supported by the National Institutes of Health, the National Science Foundation and the United States Department of Agriculture.

REFERENCES

Abramovitch, R.B. and Martin, G.B. (2004) Strategies used by bacterial pathogens to suppress plant defenses. *Curr. Opin. Plant Biol.* **7**, 356–364.

Akita, M. and Valkonen, J.P. (2002) A novel gene family in moss (*Physcomitrella patens*) shows sequence homology and a phylogenetic relationship with the TIR-NBS class of plant disease resistance genes. *J. Mol. Evol.* **55**, 595–605.

Allen, R.L., Bittner-Eddy, P.D., Grenvitte-Briggs, L.J., Meitz, J.C., Rehmany, A.P., Rose, L.E. and Beynon, J.L. (2004) Host–parasite coevolutionary conflict between Arabidopsis and downy mildew. *Science*, **306**, 1957–1960.

Ashfield, T., Ong, L.E., Nobuta, K., Schneider, C.M. and Innes, R.W.

(2004) Convergent evolution of disease resistance gene specificity in two flowering plant families. *Plant Cell*, **16**, 309–318.

Axtell, M.J. and Staskawicz, B.J. (2003) Initiation of RPS2-specified disease resistance in Arabidopsis is coupled to the AvrRpt2-directed elimination of RIN4. *Cell*, **112**, 369–377.

Bai, J., Pennill, L.A., Ning, J., Lee, S.W., Ramalingam, J., Webb, C.A., Zhao, B., Sun, Q., Nelson, J.C., Leach, J.E. and Hulbert, S.H. (2002) Diversity in nucleotide binding site-leucine-rich repeat genes in cereals. *Genome Res*. **12**, 1871–1884.

Baumgarten, A., Cannon, S., Spangler, R. and May, G. (2003) Genome-level evolution of resistance genes in *Arabidopsis thaliana*. *Genetics*, **165**, 309–319.

Belkhadir, Y., Subramaniam, R. and Dangl, J.L. (2004) Plant disease resistance protein signaling: NBS-LRR proteins and their partners. *Curr. Opin. Plant Biol*. **7**, 391–399.

Bergelson, J., Kreitman, M., Stahl, E.A. and Tian, D. (2001) Evolutionary dynamics of plant R-genes. *Science*, **292**, 2281–2285.

Botella, M.A., Parker, J.E., Frost, L.N., Bittner-Eddy, P.D., Beynon, J.L., Daniels, M.J., Holub, E.B. and Jones, J.D. (1998) Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct *Peronospora parasitica* avirulence determinants. *Plant Cell*, **10**, 1847–1860.

Brown, J.K.M. (2003) A cost of disease resistance: paradigm or peculiarity? *Trends Genet*. **19**, 667–671.

Cannon, S.B., Zhu, H., Baumgarten, A.M., Spangler, R., May, G., Cook, D.R. and Young, N.D. (2002) Diversity, distribution, and ancient taxonomic relationships within the TIR and non-TIR NBS-LRR resistance gene subfamilies. *J. Mol. Evol*. **54**, 548–562.

Chang, J.H., Goel, A.K., Grant, S.R. and Dangl, J.L. (2004) Wake of the flood: ascribing functions to the wave of type III effector proteins of phytopathogenic bacteria. *Curr. Opin. Microbiol*. **7**, 11–18.

Dangl, J.L. and Jones, J.D. (2001) Plant pathogens and integrated defence responses to infection. *Nature*, **411**, 826–833.

Deslandes, L., Olivier, J., Peeters, N., Feng, D.X., Khounlotham, M., Boucher, C., Somssich, I., Genin, S. and Marco, Y. (2003) Physical interaction between RRS1-R., a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus. *Proc. Natl Acad. Sci. USA*, **100**, 8024–8029.

Dodds, P.N., Lawrence, G.J., Catanzariti, A.M., Ayliffe, M.A. and Ellis, J.G. (2004) The *Melampsora lini* AvrL567 avirulence genes are expressed in haustoria and their products are recognized inside plant cells. *Plant Cell*, **16**, 755–768.

Dodds, P., Lawrence, G., Catanzariti, A.-M., Teh, T., Wang, C.-I.A., Ayliffe, M.B., K. and Ellis, J. (2006) Direct protein interaction underlies gene-for-gene specificity and co-evolution of the flax L5/L6/L7 resistance genes and flax rust AvrL567 avirulence genes. *Proc. Natl Acad. Sci. USA*, **103**, 8888–8893.

Dodds, P.N., Lawrence, G.J. and Ellis, J.G. (2001) Contrasting modes of evolution acting on the complex N locus for rust resistance in flax. *Plant J*. **27**, 439–453.

Ellis, J., Dodds, P. and Pryor, T. (2000) Structure, function and evolution of plant disease resistance genes. *Curr. Opin. Plant Biol*. **3**, 278–284.

Ellis, J.G., Lawrence, G.J., Luck, J.E. and Dodds, P.N. (1999) Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. *Plant Cell*, **11**, 495–506.

Filkowski, J., Yeoman, A., Kovalchuk, O. and Kovalchuk, I. (2004) Systemic plant signal triggers genome instability. *Plant J*. **38**, 1–11.

Flor, H.H. (1955) Host–parasite interactions in flax – its genetics and other implications. *Phytopathology*, **45**, 680–685.

Gopalan, S., Bauer, D.W., Alfano, J.R., Lonlello, A.O., He, S.Y. and Collmer, A. (1996) Expression of the *Pseudomonas syringae* avirulence protein AvrB in plant cells alleviates its dependence on the hypersensitive response and pathogenicity (Hrp) secretion system in eliciting genotype-specific hypersensitive cell death. *Plant Cell*, **8**, 1095–1105.

Grant, M.R., McDowell, J.M., Sharpe, A.G., de Torres Zabala, M., Lydiate, D.J. and Dangl, J.L. (1998) Independent deletions of a pathogen-resistance gene in Brassica and Arabidopsis. *Proc. Natl Acad. Sci. USA*, **95**, 15843–15848.

Greenberg, J.T. and Vinatzer, B.A. (2003) Identifying type III effectors of plant pathogens and analyzing their interaction with plant cells. *Curr. Opin. Microbiol*. **6**, 20–28.

Heidel, A.J., Clarke, J.D., Antonovics, J. and Dong, X. (2004) Fitness costs of mutations affecting the systemic acquired resistance pathway in *Arabidopsis thaliana*. *Genetics*, **168**, 2197–2206.

Heil, M. (2002) Ecological costs of induced resistance. *Current Opinion Plant Biol*. **5**, 345–350.

Henk, A.D., Warren, R.F. and Innes, R.W. (1999) A new Ac-like transposon of Arabidopsis is associated with a deletion of the RPS5 disease resistance gene. *Genetics*, **151**, 1581–1589.

Hulbert, S.H., Webb, C.A., Smith, S.M. and Sun, Q. (2001) Resistance gene complexes: evolution and utilization. *Annu. Rev. Phytopathol*. **39**, 285–312.

Innes, R.W. (2004) Guarding the goods. New insights into the central alarm system of plants. *Plant Physiol*. **135**, 695–701.

Jia, Y., McAdams, S.A., Bryan, G.T., Hershey, H.P. and Valent, B. (2000) Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *EMBO J*. **19**, 4004–4014.

Jones, D.A. and Jones, J.D.G. (1996) The roles of leucine rich repeats in plant defences. *Adv. Bot. Res. Adv. Plant Pathol*. **24**, 90–167.

Kawai, T. and Akira, S. (2005) Pathogen recognition with Toll-like receptors. *Curr. Opin. Immunol*. **17**, 338–344.

Kim, M.G., da Cunha, L., McFall, A.J., Belkhadir, Y., DebRoy, S., Dangl, J.L. and Mackey, D. (2005) Two *Pseudomonas syringae* type III effectors inhibit RIN4-regulated basal defense in Arabidopsis. *Cell*, **121**, 749–759.

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

Kobe, B. and Kajava, A.V. (2001) The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol*. **11**, 725–732.

Kooman-Gersman, M., Honée, G., Bonnema, G. and De Wit, P.J.G.M. (1996) A high-affinity binding site for the AVR9 peptide elicitor of *Cladosprium fulvum* is present on plasma membranes of tomato and other Solanaceous plants. *Plant Cell*, **8**, 929–938.

Korves, T.M. and Bergelson, J. (2003) A developmental response to pathogen infection in Arabidopsis. *Plant Physiol*. **133**, 339–347.

Korves, T. and Bergelson, J. (2004) A novel cost of R. gene resistance in the presence of disease. *Am. Natur*. **163**, 489–504.

Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M. and Hohn, B. (2003) Pathogen-induced systemic plant signal triggers DNA rearrangements. *Nature*, **423**, 760–762.

Kreitman, M. and Akashi, H. (1995) Molecular evidence for natural selection. *Annu. Rev. Ecol. Syst*. **26**, 403–422.

Kruger, J., Thomas, C.M., Golstein, C., Dixon, M.S., Smoker, M., Tang, S., Mulder, L. and Jones, J.D. (2002) A tomato cysteine protease required for Cf-2-dependent disease resistance and suppression of autonecrosis. *Science*, **296**, 744–747.

Kuang, H., Wei, F., Marano, M.R., Wirtz, U., Wang, X., Liu, J., Shum, W.P., Zaborsky, J., Tallon, L.J., Rensink, W., Lobst, S., Zhang, P.,

Tornqvist, C.E., Tek, A., Bamberg, J., Helgeson, J., Fry, W., You, F., Luo, M.C., Jiang, J., Robin Buell, C. and Baker, B. (2005) The R1 resistance gene cluster contains three groups of independently evolving, type I, R1 homologues and shows substantial structural variation among haplotypes of *Solanum demissum*. *Plant J.* **44**, 37–51.

Kuang, H., Woo, S.S., Meyers, B.C., Nevo, E. and Michelmore, R.W. (2004) Multiple genetic processes result in heterogeneous rates of evolution within the major cluster disease resistance genes in lettuce. *Plant Cell*, **16**, 2870–2894.

Leister, D. (2004) Tandem and segmental gene duplication and recombination in the evolution of plant disease resistance gene. *Trends Genet*, **20**, 116–122.

Leister, R.T., Ausubel, F.M. and Katagiri, F. (1996) Molecular recognition of pathogen attack occurs inside of plant cells in plant disease resistance specified by the *Arabidopsis* genes *RPS2* and *RPM1*. *Proc. Natl Acad. Sci. USA*, **93**, 15497–15502.

Leister, D., Kurth, J., Laurie, D.A., Yano, M., Sasaki, T., Devos, K., Graner, A. and Schulze-Lefert, P. (1998) Rapid reorganization of resistance gene homologues in cereal genomes. *Proc. Natl Acad. Sci. USA*, **95**, 370–375.

Lucht, J.M., Mauch-Mani, B., Steiner, H.Y., Metraux, J.P., Ryals, J. and Hohn, B. (2002) Pathogen stress increases somatic recombination frequency in Arabidopsis. *Nat. Genet.* **30**, 311–314.

Luck, J.E., Lawrence, G.J., Dodds, P.N., Shepherd, K.W. and Ellis, J.G. (2000) Regions outside of the leucine-rich repeats of flax rust resistance proteins play a role in specificity determination. *Plant Cell*, **12**, 1367–1377.

Luderer, R., Rivas, S., Nurnberger, T., Mattei, B., Van den Hooven, H.W., Van der Hoorn, R.A., Romeis, T., Wehrfritz, J.M., Blume, B., Nennstiel, D., Zuidema, D., Vervoort, J., De Lorenzo, G., Jones, J.D., De Wit, P.J. and Joosten, M.H. (2001) No evidence for binding between resistance gene product Cf-9 of tomato and avirulence gene product AVR9 of *Cladosporium fulvum*. *Mol. Plant–Microbe Interact.* **14**, 867–876.

Mackey, D., Belkhadir, Y., Alonso, J.M., Ecker, J.R. and Dangl, J.L. (2003) Arabidopsis RIN4 is a target of the type III virulence effector AvrRpt2 and modulates RPS2-mediated resistance. *Cell*, **112**, 379–389.

Mackey, D., Holt, B.F., Wiig, A. and Dangl, J.L. (2002) RIN4 interacts with *Pseudomonas syringae* type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell*, **108**, 743–754.

Martin, G.B., Bogdanove, A.J. and Sessa, G. (2003) Understanding the functions of plant disease resistance proteins. *Annu. Rev. Plant Biol.* **54**, 23–61.

Mauricio, R., Stahl, E.A., Korves, T., Tian, D., Kreitman, M. and Bergelson, J. (2003) Natural selection for polymorphism in the disease resistance gene *Rps2* of *Arabidopsis thaliana*. *Genetics*, **163**, 735–746.

Mayer, M.P. and Bukau, B. (1999) Molecular chaperones: the busy life of Hsp90. *Curr. Biol.* **9**, R322–R325.

McClintock, B. (1984) The significance of responses of the genome to challenge. *Science*, **226**, 792–801.

McDowell, J.M. (2004) Convergent evolution of disease resistance genes. *Trends Plant Sci.* **9**, 315–317.

McDowell, J.M., Dhandaydham, M., Long, T.A., Aarts, M.G., Goff, S., Holub, E.B. and Dangl, J.L. (1998) Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. *Plant Cell*, **10**, 1861–1874.

Meyers, B.C., Dickerman, A.W., Michelmore, R.W., Sivaramakrishnan, S., Sobral, B.W. and Young, N.D. (1999) Plant disease resistance genes encode members of an ancient and diverse protein family within the nucleotide-binding superfamily. *Plant J.* **20**, 317–332.

Meyers, B.C., Kaushik, S. and Nandety, R.S. (2005) Evolving disease resistance genes. *Curr. Opin. Plant Biol.* **8**, 129–134.

Meyers, B.C., Kozik, A., Griego, A., Kuang, H.H. and Michelmore, R.W. (2003) Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *Plant Cell*, **15**, 809–834.

Meyers, B.C., Morgante, M. and Michelmore, R.W. (2002) TIR-X and TIR-NBS proteins: two new families related to disease resistance TIR-NBS-LRR proteins encoded in Arabidopsis and other plant genomes. *Plant J.* **32**, 77–92.

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

Michelmore, R.W. and Meyers, B.C. (1998) Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process. *Genome Res.* **8**, 1113–1130.

Moffett, P., Farnham, G., Peart, J. and Baulcombe, D.C. (2002) Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death. *EMBO J.* **21**, 4511–4519.

Mondragon-Palomino, M., Meyers, B.C., Michelmore, R.W. and Gaut, B.S. (2002) Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. *Genome Res.* **12**, 1305–1315.

Monosi, B., Wisser, R.J., Pennill, L. and Hulbert, S.H. (2004) Full-genome analysis of resistance gene homologues in rice. *Theor. Appl. Genet.* **109**, 1434–1447.

Nei, M. and Rooney, A.P. (2005) Concerted and birth-and-death evolution of multigene families. *Annu. Rev. Genet.* **39**, 121–152.

Nimchuk, Z., Marois, E., Kjemtrup, S., Leister, R.T., Katagiri, F. and Dangl, J.L. (2000) Eukaryotic fatty acylation drives plasma membrane targeting and enhances function of several type III effector proteins from *Pseudomonas syringae*. *Cell*, **101**, 353–363.

Nobuta, K., Ashfield, T., Kim, S. and Innes, R.W. (2005) Diversification of non-TIR class NB-LRR genes in relation to whole-genome duplication events in Arabidopsis. *Mol. Plant–Microbe Interact.* **18**, 103–109.

Nurnberger, T., Brunner, F., Kemmerling, B. and Piater, L. (2004) Innate immunity in plants and animals: striking similarities and obvious differences. *Immunol. Rev.* **198**, 249–266.

Pan, Q., Liu, Y.S., Budai-Hadrian, O., Sela, M., Carmel-Goren, L., Zamir, D. and Fluhr, R. (2000) Comparative genetics of nucleotide binding site-leucine rich repeat resistance gene homologues in the genomes of two dicotyledons: tomato and arabidopsis. *Genetics*, **155**, 309–322.

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

Parniske, M. and Jones, J.D. (1999) Recombination between diverged clusters of the tomato Cf-9 plant disease resistance gene family. *Proc. Natl Acad. Sci. USA*, **96**, 5850–5855.

Rathjen, J.P. and Moffett, P. (2003) Early signal transduction events in specific plant disease resistance. *Curr. Opin. Plant Biol.* **6**, 300–306.

Richly, E., Kurth, J. and Leister, D. (2002) Mode of amplification and reorganization of resistance genes during recent *Arabidopsis thaliana* evolution. *Mol. Biol. Evol.* **19**, 76–84.

Rooney, H.C., Van't Klooster, J.W., van der Hoorn, R.A., Joosten, M.H., Jones, J.D. and de Wit, P.J. (2005) Cladosporium Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science*, **308**, 1783–1786.

Rose, L.E., Bittner-Eddy, P.D., Langley, C.H., Holub, E.B., Michelmore, R.W. and Beynon, J.L. (2004) The maintenance of extreme amino acid diversity at the disease resistance gene, RPP13, in *Arabidopsis thaliana*. *Genetics*, **166**, 1517–1527.

Sangster, T.A., Lindquist, S. and Queitsch, C. (2004) Under cover: causes, effects and implications of Hsp90-mediated genetic capacitance. *Bioessays*, **26**, 348–362.

Schulze-Lefert, P. (2004) Plant immunity: the origami of receptor activation. *Curr. Biol.* **14**, R22–R24.

Shao, F., Golstein, C., Ade, J., Stoutemyer, M., Dixon, J.E. and Innes, R.W. (2003) Cleavage of Arabidopsis PBS1 by a bacterial type III effector. *Science*, **301**, 1230–1233.

van't Slot, K.A.E. and Knogge, W. (2002) A dual role for microbial pathogen-derived effector proteins in plant disease and resistance. *Crit. Rev. Plant Sci.* **21**, 229–271.

Smith, S.M. and Hulbert, S.H. (2005) Recombination events generating a novel Rp1 race specificity. *Mol. Plant–Microbe Interact.* **18**, 220–228.

Smith, S.M., Pryor, A.J. and Hulbert, S.H. (2004) Allelic and haplotypic diversity at the rp1 rust resistance locus of maize. *Genetics*, **167**, 1939–1947.

Stahl, E.A. and Bishop, J.G. (2000) Plant–pathogen arms races at the molecular level. *Curr. Opin. Plant Biol.* **3**, 299–304.

Stahl, E.A., Dwyer, G., Mauricio, R., Kreitman, M. and Bergelson, J. (1999) Dynamics of disease resistance polymorphism at the Rpm1 locus of Arabidopsis. *Nature*, **400**, 667–671.

Thomas, C.M., Jones, D.A., Parniske, M., Harrison, K., Balint-Kurti, P.J., Hatzixanthis, K. and Jones, J.D.G. (1997) Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. *Plant Cell*, **9**, 2209–2224.

Tian, D., Traw, M.B., Chen, J.Q., Kreitman, M. and Bergelson, J. (2003) Fitness costs of R-gene-mediated resistance in *Arabidopsis thaliana*. *Nature*, **423**, 74–77.

Van der Biezen, E.A. and Jones, J.D.G. (1998) Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem. Sci.* **12**, 454–456.

Van der Hoorn, R.A., De Wit, P.J. and Joosten, M.H. (2002) Balancing selection favors guarding resistance proteins. *Trends Plant Sci.* **7**, 67–71.

Zhou, T., Wang, Y., Chen, J.Q., Araki, H., Jing, Z., Jiang, K., Shen, J. and Tian, D. (2004) Genome-wide identification of NBS genes in japonica rice reveals significant expansion of divergent non-TIR NBS-LRR genes. *Mol. Genet. Genomics*, **271**, 402–415.