The Plant Cell, Vol. 14, 435–450, February 2002, www.plantcell.org © 2002 American Society of Plant Biologists

Large-Scale Structure–Function Analysis of the Arabidopsis RPM1 Disease Resistance Protein

Pablo Tornero,[a] Ryon A. Chao,[a] William N. Luthin,[a] Stephen A. Goff,[b] and Jeffery L. Dangl[a,c,1]

[a] Department of Biology, University of North Carolina, Chapel Hill, North Carolina 27599-3280
[b] Torrey Mesa Research Institute, San Diego, California 92121-1125
[c] Curriculum in Genetics, University of North Carolina, Chapel Hill, North Carolina 27599-3280

The Arabidopsis *RPM1* gene confers resistance against *Pseudomonas syringae* expressing either the AvrRpm1 or the AvrB type III effector protein. We present an exhaustive genetic screen for mutants that no longer recognize *avrRpm1*. Using an inducible *avrRpm1* expression system, we identified 110 independent mutations. These mutations represent six complementation groups. None discriminates between *avrRpm1* and *avrB* recognition. We identified 95 *rpm1* alleles and present a detailed structure–function analysis of the RPM1 protein. Several *rpm1* mutants retain partial function, and we deduce that their residual activity is dependent on the level of *avrRpm1* signal. In these mutants, the hypersensitive response remains activated if the signal goes above a certain threshold. Missense mutations in *rpm1* are highly enriched in the nucleotide binding domain, suggesting that this region plays a key role either in the hypersensitive response associated with RPM1 activation or in RPM1 stability. Cluster analysis of *rpm1* alleles defines functionally important residues that are highly conserved between nucleotide binding site leucine-rich repeat R proteins and those that are unique to *RPM1*. Regions of RPM1 to which no loss-of-function alleles map may represent domains in which variation is tolerated and may contribute to the evolution of new *R* gene specificities.

INTRODUCTION

Plants can recognize certain products produced by pathogens and mount an appropriate disease resistance response. Soon after the rediscovery of Gregor Mendel's work, it became clear that single loci could confer resistance to otherwise susceptible lines if transferred by introgression (Biffin, 1905). Flor (1971) condensed this idea by postulating the existence of genes in the plant, called resistance (*R*) genes, that are able to interact genetically with genes of the pathogen, called avirulence (*Avr*) genes. If, and only if, both genes are present, a resistance response is triggered. Frequently, the resistance phenotype is accompanied by a programmed cell death called the hypersensitive response (HR) (reviewed by Morel and Dangl, 1997; Heath, 2000). It is not clear if HR is the cause or the consequence of the resistance response.

We work with *Pseudomonas syringae* pv *tomato* (*Pst*), a bacterial pathogen of Arabidopsis. This pathogen is extracellular and causes disease by "delivering" protein disease effectors into the host cell interior by means of an evolutionarily conserved type III secretion apparatus (Nimchuk et al., 2001). If the plant genotype infected carries an appropriate *R* gene, conditioning the recognition of one of the type III effectors, then disease resistance occurs. If recognition does not occur, colonization and disease ensue. We study the recognition of *avrRpm1* and *avrB* by Arabidopsis. The RPM1 protein recognizes the presence of either of these two sequence-unrelated type III effectors (Dangl et al., 1992; Bisgrove et al., 1994). The deduced RPM1 sequence predicts three main protein domains (Grant et al., 1995). The N terminus features a predicted coiled-coil (CC) domain (Pan et al., 2000). The C terminus is formed by ~14 leucine-rich repeat (LRR) sequences (Kobe and Deisenhofer, 1995). In the middle of the protein is a motif containing consensus sequences for a nucleotide binding site (NB) as well as homology between mammalian APAF-1, plant NB-LRR R, and *Caenorhabditis elegans* CED-4 proteins, which together constitute the NB-ARC domain (van der Biezen and Jones, 1998). APAF-1 and CED-4 are regulators of programmed cell death (Aravind et al., 1999). RPM1 is a peripheral plasma membrane protein (Boyes et al., 1998), and AvrB and AvrRpm1 also are trafficked to the host cell plasma membrane after in planta expression (Nimchuk et al., 2000).

Three screens for the loss of AvrRpm1 or AvrB recognition

[1] To whom correspondence should be addressed. E-mail dangl@email.unc.edu; fax 919-962-1625.
Article, publication date, and citation information can be found at www.plantcell.org/cgi/doi/10.1105/tpc.010393.

have been reported (Bisgrove et al., 1994; Century et al., 1995; Grant et al., 1995). These screens covered only a few thousand plants, were laborious, and obviously were not saturating. Only two genes, *RPM1* and *NDR1*, were identified by their loss of function. Mutant *ndr1* plants do not restrict the growth of bacterial pathogens containing *avrRpm1* or *avrB*. Thus, *NDR1* is required for *RPM1* function. Recognition of other sequence-unrelated *avr* genes, such as *avrRpt2* and *avrPphB* (Century et al., 1997), also is compromised in *NDR1* plants. Similarly, the *pbs2* mutation in Arabidopsis ecotype Columbia (Col-0) eliminates the recognition of the same set of *P. syringae avr* genes (Warren et al., 1999). Collectively, these results suggest that *NDR1* and *PBS2* are components of a signal transduction network required for the function of a subset of Arabidopsis *R* genes (Aarts et al., 1998). A variety of other Arabidopsis genes are required for the action of one or more *R* genes (Glazebrook, 2001), but these have no effect on *RPM1* function.

These results are consistent with large-scale mutational analysis of the barley *Ml-a* locus (Torp and Jørgensen, 1986; Jørgensen, 1988) and of flax resistance loci (Lawrence et al., 1993). The small number of loci identified in these screens prompted us to develop tools for a massive screen for genes required for *RPM1* function and subsequent disease resistance responses. By screening at least 1000-fold more mutagenized seedlings, we reasoned that our chances of identifying rare, nonlethal alleles of critical genes, or two simultaneous mutations (Wilhelmi and Preuss, 1996) required for RPM1 function, would be enhanced greatly. We present the results of this screen and our structure–function analysis of the RPM1 protein.

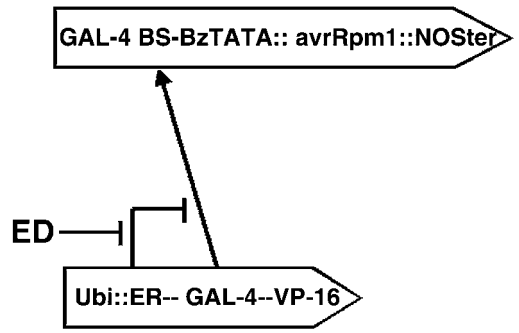

Figure 1. Scheme of the Inducible *avrRpm1* Expression System.

This system requires two T-DNAs, the "driver" (bottom) and the "inducible promoter" (top). The driver is a chimeric protein with three components: GAL-4 for specificity, estrogen receptor (ER) for induction, and VP-16 as a transcriptional activator. The inducible promoter has the GAL-4 binding sequence to increase transcription. When ED is present, transcription is activated and AvrRpm1 is produced inside the plant cell. Ubi, ubiquitin; NOSter, nopaline synthase terminator; GAL4-BS, GAL4 binding site; Bz TATA, bronze TATA box (Guyer et al., 1998).

RESULTS

An Estradiol-Inducible *avrRpm1* Expression System Retains Specificity

There are several inducible gene expression systems available for use in Arabidopsis (reviewed by Zuo and Chua, 2000). We used an inducible system that is conceptually similar to a previous system (Zuo et al., 2000) and a logical continuation of the work of another group (Guyer et al., 1998) (see Methods). Essentially, when β-estradiol (ED) is applied to the plant, a chimeric transcription factor binds a chimeric promoter that regulates *avrRpm1* expression (see Methods). Figure 1 illustrates the functional components of this system. We generated 25 independent transgenic Col-0 lines that produced a strong *avrRpm1-RPM1*–dependent response after application of ED (see below). We selected one of these at random and established a homozygous line, designated a11. We also introgressed a11 into an isogenic *rpm1-1* mutant background (Grant et al., 1995) (Table 1). This homozygous line is called a11r.

The a11 and a11r plants were essentially wild type in the absence of ED. We confirmed that these transgenic lines responded like their respective nontransgenic parents to bacterial inoculation in the absence of ED. We sometimes observed a small increase in the resistance of a11 to the virulent *Pst*DC3000 isolate after inoculation in the absence of ED. This effect was sporadic. Figure 2 illustrates the growth of *Pst*DC3000 in these lines 3 days after inoculation in one such experiment. The enhanced resistance in a11 could be attributable to leaky *avrRpm1* expression that is sufficient to trigger some defense response but not to kill the cells. We favor this explanation because the a11r line, which is *rpm1*, lacked this phenotype (Figure 2). We detected *avrRpm1* mRNA in a11 by reverse transcriptase–mediated polymerase chain reaction (RT-PCR) but not by mRNA or protein gel blot analysis (data not shown). This effect was not accompanied by detectable microscopic cell death (data not shown). The response of a11 plants to *Pst*DC3000(*avrRpm1*) was identical to that of wild-type Col-0 (*RPM1*), as shown in Figure 2. This was true as well for responses of a11 to *Pst*DC3000 containing the *avrRpt2*, *avrRps4*, and *avrPphB* genes, which were recognized in Col-0 by different NB-LRR *R* genes (data not shown). The a11 line also produced wild-type HR in response to high-dose inoculation of *Pst*DC3000 carrying any of these *avr* genes (data not shown). In sum, despite some sporadic leaky expression, the a11 and a11r lines recapitulated the normal response of Col-0 to *Pst*DC3000 alone or carrying any of several *avr* genes.

We titrated the response of soil-grown seedlings to sprayed ED, measured as *RPM1-avrRpm1*–dependent cell death. The a11 line responded to ED levels as low as 50 nM (data not shown), with a maximum response at 10 μM. The a11 plants were stunted severely, and their leaves were necrotic by 7 days after ED (10 μM; see Methods), as shown in

Table 1. *rpm1* Alleles Mentioned in This Article

Allele	Name	Mutation	Amino Acid Change	Physicochemical Distance	Conservation	Strength
Cluster 9A						
rpm1-4	(Grant et al., 1995)	3492 G to A	766 G to E	98	1	0
rpm1-7	(Grant et al., 1995)	2688 C to T	498 P to S	74	1	0
rpm1-15	E8/2	1828 C to T	211 A to V	64	1	0
rpm1-18	E17/1	2607 G to A	471 E to K	56	1	0
rpm1-19	E19/1	3492 G to A	766 G to E	98	1	0
rpm1-36	E66/3	3750 C to T	852 L to F	22	1	0
rpm1-39	E71/3	2065 C to T	290 T to I	89	1	0
rpm1-40	E72/2	2616 G to A	474 A to T	58	1	0
rpm1-41	E75/1	2170 G to A	325 G to E	98	1	0
rpm1-50	E88/12	2686 G to A	497 R to Q	43	1	0
rpm1-52	E95/1	2587 C to T	464 P to L	98	1	0
rpm1-57	E103/3	2677 C to T	494 P to L	98	1	0
rpm1-58	E108/1	2001 G to A	269 V to M	21	1	0
rpm1-72	T39/1	3966 C to T	924 L to F	22	1	0
rpm1-78	E1/11	2689 C to T	498 P to L	98	1	0
rpm1-87	E70/1	3556 G to A	787 R to H	29	1	0
rpm1-90	E92/1	1509 C to T	105 P to S	74	1	0
rpm1-93	E69/14	3667 G to A	824 G to E	98	1	1
rpm1-94	E71/6	1795 G to A	200 G to E	98	3	1
rpm1-96	E62/1	2341 C to T	382 S to F	155	1	2
rpm1-97	E67/1	3304 G to A	703 G to R	125	1	2
Cluster 9B						
rpm1-5	(Grant et al., 1995)	2346 G to A	384 G to R	125	3	0
rpm1-8	(Grant et al., 1995)	3631 A to T	812 N to I	149	1	0
rpm1-9	E1/9	2740 C to T	515 S to F	155	3	0
rpm1-10	E3/2	2595 G to A	467 G to R	125	1	0
rpm1-23	E29/2	2346 G to A	384 G to R	125	3	0
rpm1-24	E31/1	3409 C to T	738 S to F	155	1	0
rpm1-25	E32/1	2512 C to T	439 S to F	155	3	0
rpm1-29	E40/5	2346 G to A	384 G to R	125	3	0
rpm1-47	E84/4	1716 G to A	174 G to R	125	2	0
rpm1-61	E118/1	2346 G to A	384 G to R	125	3	0
rpm1-84	E46/3	1324 C to T	43 S to F	155	3	0
rpm1-85	E57/1	2346 G to A	384 G to R	125	3	0
rpm1-86	E69/3	1716 G to A	174 G to R	125	2	0
rpm1-89	E90/1	2512 C to T	439 S to F	155	3	0
Cluster 9C						
rpm1-2	(Grant et al., 1995)	2159 A to T	301 L to F	22	2	0
rpm1-6	(Grant et al., 1995)	1752 C to T	186 L to F	22	4	0
rpm1-17	E13/1	2214 G to A	340 E to K	56	3	0
rpm1-21	E21/2	2010 C to T	272 L to F	22	3	0
rpm1-26	E34/3	2332 C to T	379 A to V	64	3	0
rpm1-28	E37/1	3537 C to T	781 L to F	22	4	0
rpm1-33	E56/4	2379 G to A	395 E to K	56	3	0
rpm1-35	E62/2	1935 G to A	247 E to K	56	2	0
rpm1-37	E67/7	3054 C to T	620 P to S	74	3	0
rpm1-51	E90/3	2214 G to A	340 E to K	56	3	0
rpm1-60	E115/1	3528 C to T	778 L to F	22	3	0
rpm1-62	E119/5	1851 G to A	219 V to M	21	2	0

Continued

Table 1. (continued).

Allele	Name	Mutation	Amino Acid Change	Physicochemical Distance	Conservation	Strength
Cluster 9D						
rpm1-12	E5/3	2521 C to T	442 P to L	98	3	0
rpm1-14	E7/4	1809 G to A	205 G to R	125	5	0
rpm1-16	E10/1	1810 G to A	205 G to E	98	5	0
rpm1-27	E36/1	3055 C to T	620 P to L	98	3	0
rpm1-30	E45/7	1803 G to A	203 G to S	56	5	0
rpm1-32	E55/2	2226 C to T	344 L to F	22	5	0
rpm1-43	E79/4	1804 G to A	203 G to D	94	5	0
rpm1-44	E80/1	2134 C to T	313 T to I	89	3	0
rpm1-53	E99/3	2218 C to T	341 A to V	64	4	0
rpm1-80	E8/3	2116 G to A	307 G to E	98	3	0
Stops						
rpm1-1	(Grant et al., 1995)	3648 A deleted	818 R to F.S.[a]		NA[b]	0
rpm1-3	(Grant et al., 1995)	1157 T to A	87 Y to stop		1	0
rpm1-11	E4/3	1880 G to A	228 W to stop		4	0
rpm1-22	E22/4	3617 G to A	807 W to stop		1	0
rpm1-31	E53/1	2078 G to A	294 W to stop		1	0
rpm1-38	E68/1	2025 C to T	277 Q to stop		2	0
rpm1-42	E76/4	3239 G to A	681 W to stop		1	0
rpm1-59	E111/1	1845 C to T	217 Q to stop		1	0
rpm1-73	T47/4	2583-2861 deleted	463 to 554 deleted		NA	0
rpm1-79	E2/1	2274 C to T	360 Q to stop		1	0
rpm1-81	E33/3	3617 G to A	807 W to stop		1	0
rpm1-83	E18/6	3617 G to A	807 W to stop		1	0
rpm1-91	E112/1	2384 G to A	396 W to stop		5	0
No sequence						Strength
rpm1-13, 20, 34, 45, 46, 48, 49, 54, 55, 56, 63 to 71, 74 to 77, 82, 88, 92						0
rpm1-95, 98 to 103						2

"Name" indicates the reference in the case of *rpm1-1* to *rpm1-8* or the original mutant designation from our screen. "E" indicates an EMS mutation and "T" indicates a T-DNA mutation. The number indicates the M2 family in which each was found. "Mutation" is the nucleotide change, considering that the first nucleotide of the first codon of the open reading frame is 1196 and the stop codon is 3972 (which correspond to ATG 68117 and stop 65337 in the genomic bacterial artificial chromosome F17A9). "Amino Acid Change" is the actual amino acid mutation. "Physicochemical Distance" reflects the nature of the amino acid substitution (determined by the physicochemical distance). In the column "Conservation," we determined the degree of conservation of the wild-type amino acid at that position determined by alignment of 36 CC-NBS-LRR–predicted genes from Arabidopsis; 5 indicates positions 100% identical, 4 indicates 100% conserved, 3 indicates 75 to 99% conserved, 2 indicates 50 to 74% conserved, and 1 indicates <50%. "Strength" is an indication of the remaining wild-type function; loss of function is assigned a 1, intermediate with no HR is assigned a 2, and intermediate with HR is assigned a 3. To preserve the maximum information, we present this table by clusters, as presented in Figure 9.

[a] F.S., frameshift.

[b] NA, not available.

Figures 3A and 3B. It is important that ED did not spread systemically (i.e., the newly emerging leaves were not affected) and that ED-treated a11 plants survived. In contrast, a11r plants grew unaffected, as shown in Figures 3A and 3B. This phenotype reflects the expected mutant phenotype we hoped to recover using a11 as the parental line for mutagenesis. We performed RT-PCR after ED application on a11 and a11r and observed a clear induction of *avrRpm1* mRNA in both lines (data not shown). We used trypan blue staining to reveal dead cells early in the post-ED time course (Koch and Slusarenko, 1990). There was strong cell death on treated a11 leaves by 1 day after ED spray (Figure 3C). When 10 μM ED was hand infiltrated into a11 leaves, macroscopic cell death response mirrored the HR produced by *Pst*DC3000(*avrRpm1*) on Col-0 in both phenotype and timing (5 to 7 hr after inoculation; data not shown). Importantly, both the cell death and macroscopic symptoms associated with ED induction of *avrRpm1* expression were *RPM1* dependent, because a11r plants sprayed with ED did not exhibit any symptoms (Figure 3D), even after hand infiltration of ED (data not shown). Thus, our ED-inducible AvrRpm1 expression system recapitulated at least the cell death features of the incompatible *RPM1-avrRpm1* interaction.

Mutagenesis and Screening for Loss of *avrRpm1*-Dependent Phenotypes

We mutagenized the a11 line with ethyl methanesulfonate (EMS) and T-DNA, as described in Methods. We sprayed 10 μM ED onto mutagenized M2 generation plants at 7 and 14 days after germination and scored for the loss of *RPM1*-dependent response 21 days after germination. We screened ~3600 plants per family (172 M2 families) to ensure the recovery of all mutations with a 95% probability (Malmberg, 1993). Therefore, we screened $>5 \times 10^5$ M2 seedlings. We recovered mutants from 151 families. Mutants that did not respond to ED were isolated, allowed to self, and tested in the M3 generation for loss of response to high doses of *Pst*DC3000(*avrRpm1*) to exclude mutations in the transgenic expression system. We confirmed 110 mutants and defined two phenotypic classes after *Pst*DC3000 (*avrRpm1*) inoculation. The first set of 101 mutants exhibited no HR (class I mutants). We also found nine mutants that produced some response to high doses of *Pst*DC3000 (*avrRpm1*), although they clearly did not respond to ED (class II mutants; described below). We also found 31 mutants that were still fully resistant to *Pst*DC3000(*avrRpm1*) upon retesting, although they did not respond to ED. We sequenced the *avrRpm1* transgenes from these 31 mutants and found no mutations. Hence, these were probable mutations in the inducible system and were not analyzed further.

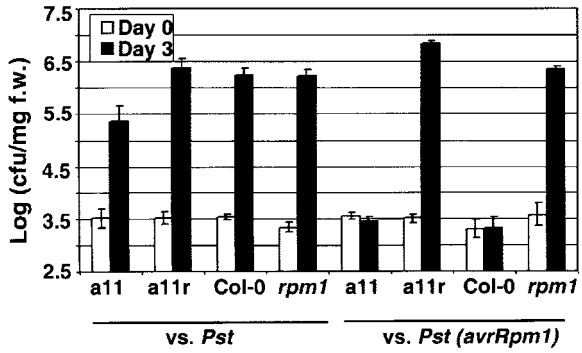

Figure 2. The Inducible *avrRpm1* Expression System Does Not Affect the Response of the Plants to Pathogens.

Plants carrying a transgenic copy of *avrRpm1* under the control of the inducible expression system in a Col-0 background (a11) or the same transgenes introgressed into the *rpm1-1* background (a11r), together with their corresponding untransformed backgrounds (Col-0 and *rpm1-1*, respectively), were inoculated by dipping with *Pst*DC3000 (*Pst*) or *Pst*DC3000(*avrRpm1*) [*Pst*(*avrRpm1*)]. At time 0 (open columns) and day 3 (closed columns), bacteria were extracted from the plants and measured (Tornero and Dangl, 2001). Bacterial numbers are expressed as log of colony-forming units (cfu) per milligram fresh weight (f.w.).
Error bars indicate ±SE.

We crossed all 110 confirmed mutants to a11r to identify *rpm1* alleles. Mutants that complemented a11r in the F1 and F2 generations for resistance to *Pst*DC3000(*avrRpm1*) were intercrossed to determine the number of complementation groups. We identified six complementation groups, including 95 alleles of *rpm1*. We named the genes defined by the remaining complementation groups *LRA*, for loss of recognition to *AvrRpm1*. Because *PBS2* is required for *RPM1* function (see Introduction), we performed complementation analysis with *pbs2-1* (Warren et al., 1999). *lra1* (eight alleles) is *pbs2* (P. Tornero and J.L. Dangl, unpublished data). We recovered four independent *lra2* alleles, but only one allele each of *lra4*, *lra5*, and *lra6*. These will be described elsewhere (P. Tornero and J.L. Dangl, unpublished data). None of our mutants discriminated *Pst*DC3000(*avrRpm1*) from *Pst*DC3000(*avrB*).

We found mutations in 62 of the 95 *rpm1* alleles (60 class I and 2 class II alleles). The 33 unsequenced alleles included those that did not give a PCR product on the first attempt (18 of 33 from EMS and 15 of 33 from T-DNA). Thus, nearly all of the T-DNA mutations produced *rpm1* alleles that could not be amplified by PCR, because the T-DNA either integrated in the open reading frame (ORF) or caused a deletion. We also did not pursue alleles in which the mutation was not found in the ORF on the first attempt (which could represent promoter mutations).

Weak and Intermediate *rpm1* Alleles Discriminate Different Doses of AvrRpm1

We found 11 intermediate or weak alleles, defined by the strength of disease symptoms and in planta growth of *Pst*DC3000(*avrRpm1*) (Tornero and Dangl, 2001). Four of them were sequenced, and we identified missense mutations in each. These mutations did not cluster (data not shown). Although these *rpm1* alleles did not exhibit any macroscopic phenotype after ED spray, nine of them were class II mutants as defined above. These responded with a macroscopic HR when inoculated with high levels of *Pst*DC3000(*avrRpm1*) (alleles *rpm1-95* to *rpm1-103*; Table 1). The two intermediate alleles that did not produce HR were *rpm1-93* and *rpm1-94*. We observed some microscopic trypan blue staining after ED application in the intermediate alleles *rpm1-93* and *rpm1-95*, as seen in Figures 4A and 4B, respectively. This light staining did not correspond to complete cell collapse. Figure 4C shows the growth of *Pst*DC3000(*avrRpm1*) 3 days after inoculation in the four sequenced intermediate alleles. Similar results were obtained with *Pst*DC3000(*avrB*) (data not shown). We additionally measured the growth of *Pst*DC3000(*avrRpt2*) (Innes et al., 1993) and *Pst*DC3000(*avrPphB*) (Simonich and Innes, 1995) and observed wild-type resistance in each case. The growth of compatible *Pst*DC3000 in these four *rpm1* alleles was comparable to that in the parental line (Figure 4D). This finding proves that the weak loss of *rpm1* function alleles was not the combinatorial outcome of a loss-of-function *rpm1*

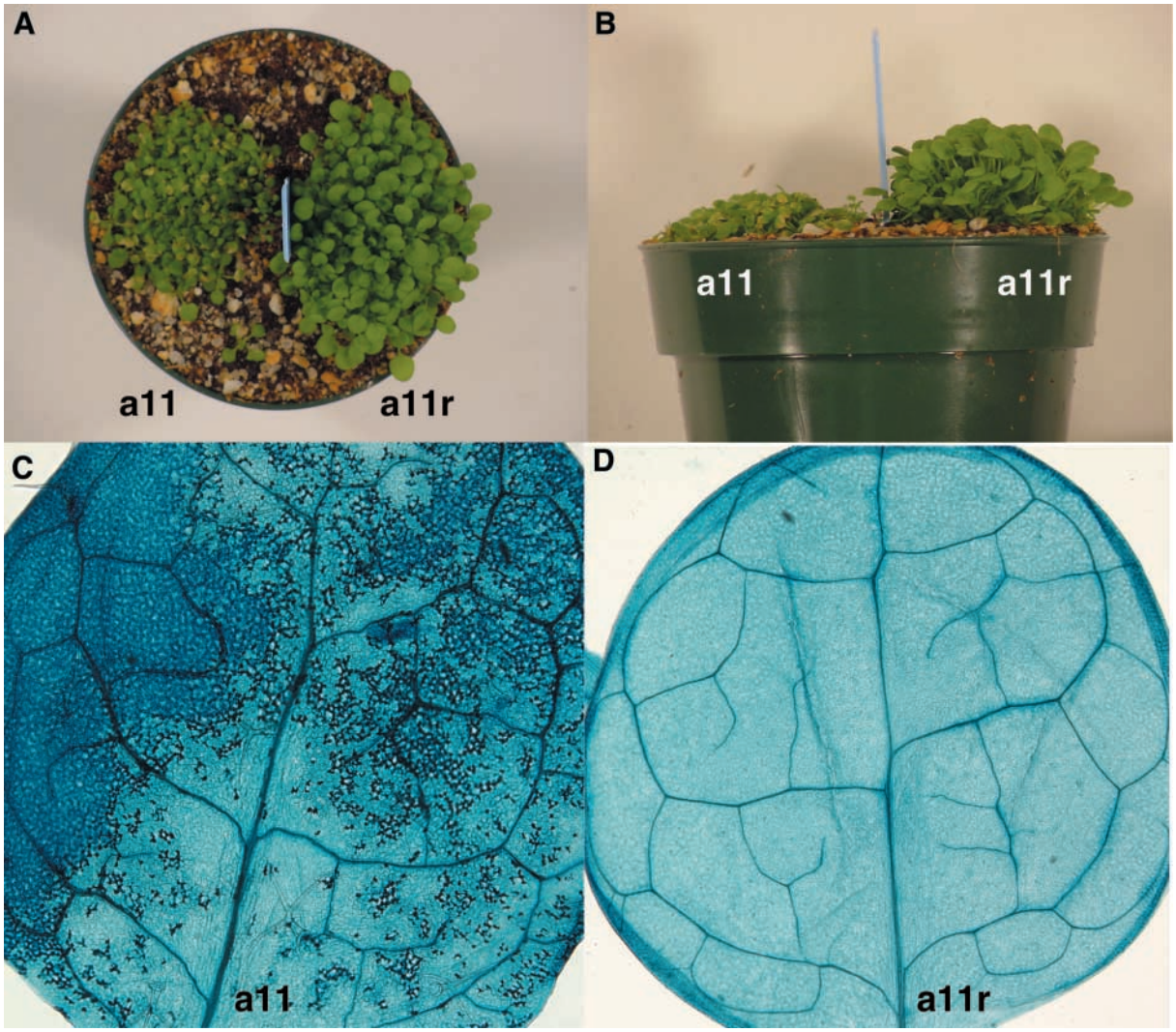

Figure 3. *RPM1*-Dependent Response after Inducible *avrRpm1* Expression.

(A) Macroscopic effect of 10 μM ED spray in lines a11 (left) and a11r (right) 5 days after spray.
(B) Similar to **(A)** but viewed from the side of the pot.
(C) Microscopic effect of the same treatment 1 day after 10 μM ED spray in line a11. Leaves of treated plants were stained with trypan blue to reveal the pattern of the veins and the HR.
(D) Similar to **(C)** but in line a11r. In this line, there is no cell death.

allele with an enhanced disease resistance mutation at a second site.

rpm1 Loss-of-Function Mutations Are Not Distributed Randomly

The remaining 84 *rpm1* alleles were all full loss of function. They did not respond to ED, they exhibited no HR after high-dose *Pst*DC3000(*avrRpm1*) inoculation, and they were visibly susceptible to low levels of these bacteria. Figure 5 illustrates the growth of *Pst*DC3000(*avrRpm1*) 3 days after inoculation into several of these strong loss-of-function alleles. They also were all susceptible to *Pst*DC3000(*avrB*) (data not shown). We sequenced 58 of these 84 *rpm1* alleles. The 11 premature termination alleles (from both our screen and the previous screen; Grant et al., 1995) did not exhibit preferential localization (data not shown). We considered mutations independent if they came from different families of M2 seed and thus recovered identical mutations indepen-

dently from different seed families (Table 1) summed over our screen and the previous screen (Bisgrove et al., 1994; Grant et al., 1995). We found that the polar uncharged amino acids are more frequent targets than we expected from a random distribution (determined by χ^2 test with $\alpha = 0.05$), whereas charged amino acids are less frequent targets. This finding could mean that charged amino acids are in general less important for the recognition of AvrRpm1 and subsequent signaling or that alteration of polar amino acids has a higher likelihood of destabilizing the RPM1 protein. On the other hand, no particular group of amino acids was created preferentially by mutation (Nivard et al., 1999) (Table 2).

We next analyzed the distribution of 47 missense loss-of-function *rpm1* alleles together with the six additional missense *rpm1* alleles described previously (Bisgrove et al., 1994; Grant et al., 1995). We did not include in this analysis alleles with frameshifts, internal deletions, or premature stops (Table 1). We plotted the frequency of missense mutations in windows of 50 amino acids along the length of RPM1 and noted a striking enrichment of mutations in, and C-terminal to, the NB-ARC domain, as displayed in Figure 6A. We found 36 of 53 mutations between amino acids 200 and 500. Thus, this region has 69% of the mutations in only 31% of the protein. Figure 6B illustrates the inverse of the statistical probability of this distribution being random assuming

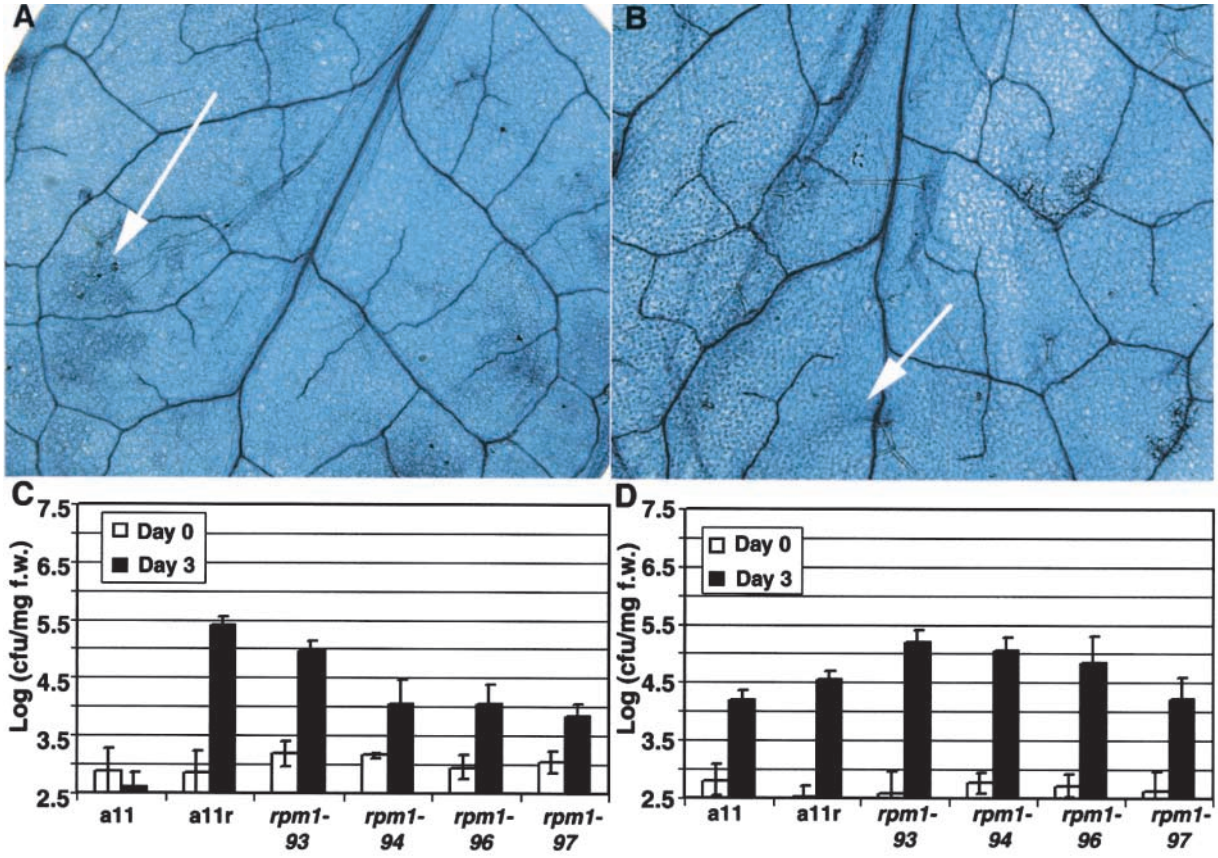

Figure 4. Characterization of Intermediate *rpm1* Alleles.

(A) The *rpm1-93* allele was treated with 10 μM ED and stained with trypan blue 1 day later. Note that the staining, although darker than in Figure 3D, is not as strong as in Figure 3C. White arrows here and in **(B)** point to a dark area of staining without cell death.

(B) The same experimental regimen using the *rpm1-95* allele.

(C) Growth of *Pst*DC3000(*avrRpm1*) in several intermediate alleles and the corresponding controls. These intermediate alleles are incompletely compromised in *RPM1* function. Bacterial numbers here and in **(D)** are expressed as log of colony-forming units (cfu) per milligram fresh weight (f.w.).

(D) Growth of *Pst*DC3000 in the same set of alleles as in **(C)**. The intermediate alleles allow the growth of virulent bacteria to levels equal to, or higher than, the corresponding controls.

Error bars indicate ±SE.

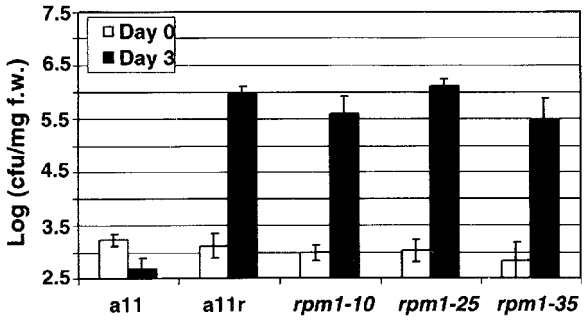

Figure 5. Characterization of Loss-of-Function *rpm1* Alleles.

Growth of *Pst*DC3000(*avrRpm1*) in three loss-of-function alleles and the corresponding controls. These loss-of-function alleles were chosen randomly and allow the same growth of *Pst*DC3000(*avrRpm1*) as observed in a11r. Bacterial numbers are expressed as log of colony-forming units (cfu) per milligram fresh weight (f.w.).

that, if random, the distribution would follow a Poisson distribution (Chou, 1969). Because we have 53 mutations in a total of 19 50–amino acid windows, the average is 2.8 mutations per 50–amino acid window. We calculated the probability of the observed distribution being random in a particular 50–amino acid window. The high q ($q = 1 - p$) values in Figure 6B indicate 50–amino acid windows with significant deviations from random distribution. For instance, the inverse of the probability of the distribution being random if we found eight missense mutations in a 50–amino acid window is 99.4%. Note that the probability also is high (93.9%) in those cases in which no mutations were found in a particular window (positions 50 to 100, 550 to 600, and 650 to 700), because lack of mutation in those windows also contradicts random distribution.

The statistical significance of our distribution was calculated under the assumption that all nucleotides in *RPM1* are equally susceptible to mutagenesis. Yet, because EMS preferentially produces a G-to-A mutation (and therefore changes a G-C base pair to an A-T base pair) (Sega, 1984), not all mutations are possible. Depending on local G+C content along a gene, it could be that in a given 50–amino acid window, 150 mutations are possible, whereas in another 50–amino acid window, no mutation is possible. In *RPM1*, this is not the case. We identified all of the possible EMS missense and nonsense mutation sites in *RPM1* and found no preferential distribution (average of 38.8 ± 5.2 possible mutations per 50 amino acids). Therefore, the deviation from random distribution described above has clear significance.

The results depicted in Figure 7A indicate that *rpm1* missense mutations did not necessarily target the highly conserved amino acid residues in the NB-ARC domain. For example, although there were 13 of 36 mutations in or close to two very conserved NB-ARC motifs (P-loop and GLPL), there was no clear predilection for clustering of mutations in the other motifs identified by conservation. Note the striking number of mutations in and after the RNBS-D motif as defined by Meyers et al. (1999). Again, the observed distribution of missense mutations did not mirror the distribution of the maximal number of possible EMS missense mutations in this domain (average of 7.8 ± 2.2 mutations per 10 amino acids).

Although the LRR region has been correlated with *R* gene specificity (reviewed by Ellis et al., 2000), it was not a frequent site of *rpm1* loss-of-function mutations. Figure 7B shows a manual alignment of the LRR domain representing the mutations with the backbone of leucines. For this purpose, we considered that the amino acids methionine, valine, and isoleucine could substitute for leucine. Eleven mutations occurred in the LRRs. Four of these are in leucine residues, and two independent mutations were in the same proline residue, which also presumably forms part of the LRR backbone (Kobe and Deisenhofer, 1994). Thus, even though the conserved hydrophobic residues and prolines constitute only 27% of the LRR (106 of 389 amino acids), they were targets for 55% of the mutations in this domain. Therefore, the missense *rpm1* alleles in the LRR were biased toward amino acids in the putative backbone of the LRR domain. These mutations might disrupt the overall LRR structure and hence its function. Additionally, they might render the RPM1 protein particularly unstable.

DISCUSSION

A Conditional Gene Expression System for Large-Scale Mutagenesis

We developed transgenic plants that allowed us to conditionally express *avrRpm1* in different backgrounds (Figure 1). In wild-type plants (line a11), we observed sporadic defense activation (Figure 2), presumably resulting from the sporadic leakiness of *avrRpm1* expression. In *rpm1* mutant plants (line a11r), this leaky expression was insufficient to produce any defense response. Therefore, the enhanced resistance that sometimes occurs in the a11 line is not caused by the conditional expression system per se (because the same transgenes are in a11r) (Kang et al., 1999); rather, it results from the fact that *RPM1* is able to recognize a vanishingly small amount of AvrRpm1 protein. We were able to detect *avrRpm1* mRNA only with RT-PCR (but not with RNA or protein gel blot analysis; data not shown). Because we did not detect cell death in unchallenged a11 plants (data not shown), recognition of very low levels of AvrRpm1 protein presumably can take place without cell death and can lead to increased resistance to virulent *Pst*DC3000. When these lines were inoculated with *Pst*DC3000(*avrRpm1*), the response was identical to that in their corresponding untransformed backgrounds (Figure 2). Thus, our inducible expression system essentially recapitulated the normal interaction of *Pst*DC3000(*avrRpm1*) and *RPM1*.

Table 2. Distribution of the *rpm1* Missense Alleles by Amino Acid Class

Class	Nonpolar	Polar	Polar Negative	Polar Positive	Stops	Total	Total (%)	Total in Coding Region (%)
Nonpolar	18	4	0	1	7	30	46.9	42.7
Polar	8	1	6	9	4	28	43.8	30.4
Polar negative	0	0	0	4	0	4	6.3	11.8
Polar positive	0	1	0	1	0	2	3.1	15.2
Total	26	6	6	15	11	64	100	100.0
Total (%)	40.6	9.4	9.4	23.4	17.2	100	100	100.0

All of the loss-of-function mutations were classified by the wild-type amino acid group (rows) or the substituted amino acid (columns). Nonpolar R groups include Ala, Val, Leu, Ile, Pro, Phe, Trp, and Met. Polar R groups include Gly, Ser, Thr, Cys, Tyr, Asn, and Gln. Negatively charged polar R groups include Asp and Glu. Positively charged polar R groups include Lys, Arg, and His.

We recovered mutations in six complementation groups. No other genetic screen for loss-of-resistance mutants in plants is as comprehensive as ours. The fact that we still recovered single alleles of three loci implies that these genes might have additional, essential functions for plant growth. If so, we presume that we recovered alleles that altered *RPM1* function but did not eliminate the putative normal function in plant growth. That *RPM1* function may require proteins with additional roles in normal plant growth is consistent with the guard hypothesis (van der Biezen and Jones, 1998; Dangl and Jones, 2001). Essentially, this hypothesis postulates that the Avr proteins normally function as virulence factors via alteration of a host component (the "guardee"). R proteins evolved to recognize the alteration of the guardee, or the complex of Avr protein and guardee, and to respond by triggering resistance (thus acting as a "guard"). One might expect that full loss-of-function mutations in guardees might be lethal. Thus, it is reasonable to think that one or more of the *LRA* genes are guardees. Alternatively, it is possible that genes defined by *lra3*, *lra4*, and *lra5* are redundant for *RPM1* function and that we recovered rare dominant negative mutations.

Note that we did not recover any *ndr1* alleles. Although *NDR1* is required for *RPM1*-mediated resistance against *Pst*DC3000(*avrRpm1*), it is not required for *RPM1*-mediated HR (Century et al., 1995). Because our screen is, in essence, for HR-negative plants, we predicted that no *ndr1* alleles would be found. To confirm this, we crossed a11 with *ndr1-1* and identified a double homozygote line. This line responded to ED-induced *avrRpm1* expression with cell death, as predicted (data not shown). We also did not recover mutants that were HR negative but then retained *RPM1* function when retested using inoculation of bacteria. Thus, we did not recover mutations in the cell death machinery per se or in loci such as the *DEFENSE NO DEATH* genes described previously (Yu et al., 1998). It is possible that the HR machinery is partially redundant. Alternatively, it is possible that the HR machinery is required for normal plant development and that mutations in these genes are lethal. We also did not find any mutant that separates *avrRpm1* from *avrB* recognition. Certainly, it is possible that an inverse mutation could be found, namely, a mutant that does not recognize *avrB* but recognizes *avrRpm1*. AvrB and AvrRpm1 share a lipid modification that targets them to the plasma membrane (Nimchuk et al., 2000). It is possible that these two Avr proteins have a similar tertiary structure or that the proteins with which they, and RPM1, interact in a potential protein complex are not mutable.

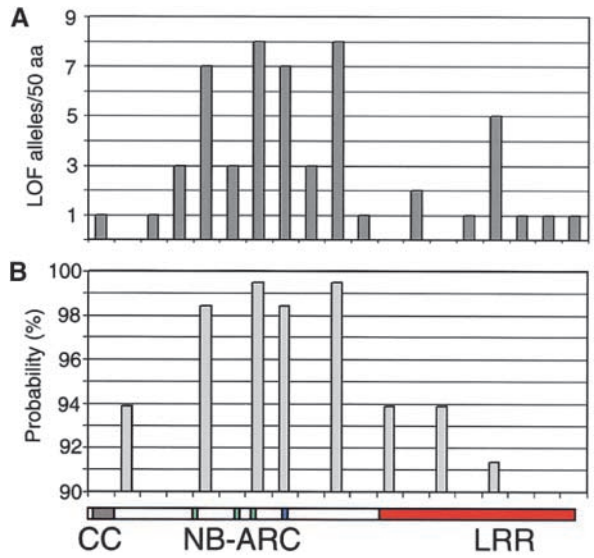

Figure 6. Loss-of-Function *rpm1* Missense Alleles Are Not Distributed Randomly.

(A) The ORF of RPM1 was divided into 50–amino acid (aa) windows, and the number of missense mutations in each window was plotted against the linear structure of RPM1. LOF, loss-of-function.

(B) Inverse of the probability that this distribution is random. For each window, we calculated the inverse of the probability that the number of mutations found in that window was random. We represent only q values of 90 to 100% ($q = 1 - p$). The scheme at bottom shows RPM1 with the CC motif (gray), nucleotide binding (NB-ARC; green for the canonical motifs P-loop, kinase 2, and kinase 3a [also called RNBS-B], and blue for the GLPL motif), and the LRR (red).

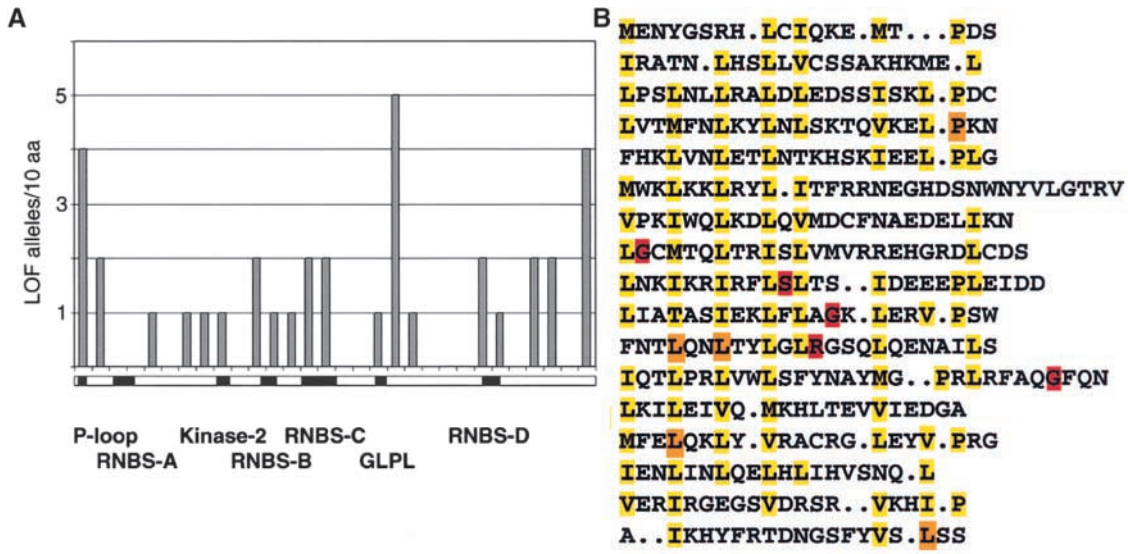

Figure 7. Loss-of-Function *rpm1* Missense Alleles in the NB-ARC and LRR Domains.

(A) Distribution of the missense mutations in the NB-ARC domain. The representation of the mutations is similar to that in Figure 6A but uses a 10–amino acid (aa) window. The scheme at bottom shows a linear representation of the NB with the motifs defined in the literature. LOF, loss-of-function.

(B) Alignment of the LRR. The core of prolines and leucines (or equivalent amino acids) is highlighted in yellow, and the mutations found are highlighted in red. If the mutations coincide with the core, the color becomes orange.

Spectrum of *rpm1* Alleles and NB-LRR Function

Of the 95 *RPM1* alleles we identified, 11 displayed some intermediate or weak resistance when challenged with *Pst*DC3000(*avrRpm1*) (Figure 4C) or *avrB* (data not shown). Most of these (nine of 11) still produced a visible HR when challenged with high levels of bacteria (data not shown). The phenotypes of these intermediate alleles suggest that the resistance response (measured as growth of bacteria) and the HR can distinguish between different levels of AvrRpm1-dependent signal. If the AvrRpm1 signal was very low, as in the absence of ED, we observed sporadic resistance to *Pst*DC3000 but no cell death. Because the intermediate *rpm1* alleles did not respond to ED with cell death (Figures 4A and 4B) yet maintained partial resistance to *Pst*DC3000(*avrRpm1*), the level of signal triggered by ED must be lower than the level of signal produced by bacteria in these low-dose inoculations. The macroscopic HR, generated by the highest dose of bacteria, would provide the highest level of signal. Accordingly, most of the *RPM1* intermediate alleles maintained this response. The idea that there might be different thresholds for HR and resistance is supported by experiments with the potato *Rx* gene. This NB-LRR protein confers resistance to the virus PVX (Bendahmane et al., 1999). There normally is no HR associated with *Rx* function, but a strong HR was triggered when a transient system was used to express the PVX coat protein, which is the corresponding Avr protein. In this case, defense was triggered first and was epistatic to HR. Thus, HR might be the upper limit outcome of increased signal levels.

By far, loss-of-function *rpm1* alleles were the most abundant mutation in our screen. This finding is consistent with smaller scale mutagenesis experiments for disease resistance in Arabidopsis (Bisgrove et al., 1994; Century et al., 1995; Grant et al., 1995), tomato (Salmeron et al., 1994), and barley (Torp and Jørgensen, 1986; Jørgensen, 1988). Before our work on *RPM1*, the largest *R* gene mutagenesis was seen in the barley *Ml-a12* gene, in which Jørgensen and colleagues isolated >20 loss-of-function *ml-a12* alleles and only three additional mutants in two other loci, later defined as *RAR1* and *RAR2* (Jørgensen, 1988; Freialdnhoven et al., 1994). Thus, our ratio of 95 *rpm1* alleles to 15 other mutations was not unexpected. A screen similar to ours was performed recently using the conditional expression of *avrRpt2* for mutagenesis of the *RPS2* response in Arabidopsis (Axtell et al., 2001). In that study, 17 new *rps2* alleles were found, and 10 of them were sequenced. Besides *rps2* alleles, only mutations in the *avrRpt2* transgene were found. The number of plants screened by these authors was in the same order of magnitude as our screen, but the conditional expression system they used was different (McNellis et al., 1998). *avrRpt2* expression was driven by the application of dexamethasone on agar plates. This chronic application of a chemical inducer known to move systemically in Arabidopsis (Aoyama and Chua, 1997) may have favored the isolation of only very strong loss-of-function mutations. In our case,

ED is not systemic in plants (data not shown), and plants that responded survived. This probably allowed us to recover weak and intermediate phenotypes.

For the purposes of analyzing the *rpm1* mutations, we also included the loss-of-function alleles described previously (Bisgrove et al., 1994; Grant et al., 1995). Pooling together the results of these screens, we found 11 stop codons that could result in a truncated RPM1 protein and two internal deletions (Table 1). If all possible EMS mutations in *RPM1* that gave rise to missense or nonsense alleles were generated, ~10% would be premature stops. Yet, the number of nonsense mutations found is higher (16% of the total) and in the range of previous reports (11% for the *hrpt* gene [Cole and Skopek, 1994] and 13% for the *vermillion* gene [Nivard et al., 1999]). Because we recovered several nonsense mutations near the C terminus of RPM1 (Table 1), we infer that nearly all possible truncations produce loss of function and therefore would be recovered. Recovery of a fraction of premature stop mutations nearer the theoretical minimum would have indicated that relatively more of the possible missense mutations in *RPM1* had affected its function. Because we recovered more than the theoretical minimum of premature stop mutations, we suggest that *RPM1* is able to sustain a considerable number of missense mutations that do not affect its function. Furthermore, there are domains of RPM1 in which missense mutations must have occurred but that did not yield mutants recoverable in our screen. No missense alleles were found in the 50–amino acid windows from residues 50 to 100, 550 to 600, and 650 to 700. The probability that this was caused by chance is low (Figure 6B). We speculate that mutations in these regions of RPM1 do not alter its function with respect to recognition of *avrRpm1* or *avrB*. If this is true, these mutations could add new specificities to RPM1. Thus, variability in these regions may help to diversify R function.

We are unable to address which of our 53 missense mutations might cause RPM1 protein instability. Thus, amino acids defined in this screen as relevant for the function of *RPM1* could be important for RPM1 stability. Recent work analyzing the structure–function correlates of other NB-LRR proteins also has this limitation (Dinesh-Kumar et al., 2000; Axtell et al., 2001). However, Tao et al. (2000) showed that their loss-of-function *rps2* alleles sometimes were even more stable than those of the wild-type protein. A similar effect was found in mutations of the *RAS* genes, with which the *R* genes have limited homology in the NB-ARC domain. Thus, certain RAS loss-of-function mutations can stabilize the protein (Feig and Cooper, 1988; Farnsworth and Feig, 1991; Chen et al., 1994).

We took advantage of the sequence alignment of 36 predicted Arabidopsis CC-NB-LRR proteins (Arabidopsis Genome Initiative, 2000). This allowed us to determine if the wild-type amino acids that were mutated were particularly conserved among this representative sample of deduced Arabidopsis NB-LRR R proteins. We found mutations in 5% of the total RPM1 amino acids that are >50% conserved (31 mutations in 620 amino acids), 6% in amino acids conserved between 50 and 75% (six mutations in 94 amino acids), 13% in amino acids conserved between 75 and 99% (21 mutations in 162 amino acids), 11% in amino acids 100% conserved (four mutations in 36 amino acids), and 43% in amino acids identical across all 36 CC-NB-LRR proteins (six mutations in 14 amino acids) (Table 1). We estimated the severity of each missense mutation by measuring the physicochemical distance (PHCD) between the wild-type and mutant amino acids (Grantham, 1974) (see Methods). The average PHCD of all of the *rpm1* alleles in this analysis was 84.7, below the generic mean of 100 (Grantham, 1974). As expected, if a particular amino acid is conserved absolutely, any mutation will produce a phenotypic effect. In contrast, the *rpm1* mutants in amino acid positions that are >50% conserved give an average PHCD of 84.6, nearly identical to the overall average. Because average changes in these nonconserved amino acids eliminate function, we propose that they are as important for RPM1 activity as the most conserved amino acids.

There is an obvious enrichment of *rpm1* missense mutations in the NB-ARC region between amino acids 200 and 500. In this region, the average 50–amino acid window has six mutations, as opposed to 1.3 mutations per 50 amino acids across the rest of the protein. Meyers et al. (1999) defined subregions in the NB-ARC based on their conservation from an alignment of 248 predicted NB-LRR *R* genes from different plant species. The P-loop is very conserved in the NB-ARC domain (Aravind et al., 1999) and sustained four mutations in our screens. The same number of mutations occurred in the conserved RNBS-C domain (Meyers et al., 1999). We found three mutants in RNBS-D, a motif specific to CC-NBS-LRR proteins (Meyers et al., 1999; Pan et al., 2000). Importantly, however, most of the *rpm1* mutations are outside of motifs clearly defined by homology. Thus, of the 36 *rpm1* alleles between amino acids 200 and 500, 14 mutations are localized in conserved motifs and 22 are not. We speculate that these 22 positions affect the function of *RPM1* by defining its specificity.

There have been two reports of directed mutagenesis of *R* genes. Dinesh-Kumar et al. (2000) constructed a series of mutants in the tobacco *N* gene. In *N*, a G219D (P-loop) mutation produced a loss of function, and two independent mutations in the corresponding *RPM1* position, G203D (*rpm1-43*) and G203S (*rpm1-30*), also eliminated function (Figure 8). This G residue is structurally relevant for nucleotide binding in RAS (Lacal, 1993). The *N* gene G216E mutation corresponds to *RPM1* G200E (*rpm1-94*). Although in *N* this mutation produces a dominant negative loss of function, the result in *RPM1* is an intermediate allele. We did not identify dominant negative alleles in our screen. Note that if a dominant negative allele was generated in our mutagenesis, it could have been detected as either a homozygote or a heterozygote and hence would have been observed more frequently in our primary screen than simple recessive loss-of-function mutations (see Methods). We did not see any

```
RPM1    (200) GMGGSGKTTL (294) WREISIALPD
RPS2    (183) GPGGVGKTTL (270) LEKTGVPRPD
N       (216) GMGGVGKTTI (309) YLEYLAGDLD
APAF-1 (147) GMAGCGKSVL (251) LKAF-----D

RPM1    (433) CFLYCSLFPVN (497) R-------PKA-------
RPS2    (407) CFLYCALFPEE (497) PSMGHTEAPKAENWRQAL
N       (443) CFLRGEEKDYI (528) MAMEAIWVSSYSSTLRFS
APAF-1 (377) EDIKDYYTDLS (467) YHQPHTLSPDQEDCMYWY
```

Figure 8. Comparison of Selected Regions of the NB across Several Genes.

RPM1, RPS2, N, and APAF-1 were aligned, and the regions of the NB mutations found in other screens are represented.

mutation in the conserved lysine of the P-loop (K222 in *N*, K206 in *RPM1*) because EMS does not affect this codon. The other directed mutagenesis study was performed by Tao et al. (2000) in the Arabidopsis gene *RPS2*. We did not find any mutations corresponding to those constructed by Tao and colleagues in the *RPS2* P-loop. *RPS2* also was the target of in planta mutagenesis. Using a system very similar to the one described here, Axtell et al. (2001) found six missense mutations in *RPS2*, among other alleles. We did not find the corresponding mutations in RPM1 for *RPS2* P276L (*RPM1* A300) or *RPS2* A456T (*RPM1* Q498), but we did find a mutation similar to *RPS2* A412V (*RPM1* S439F; Figure 8).

We did not find an enrichment of mutations in the RPM1 LRR region (Figure 7B). There is compelling evidence that the LRRs of NB-LRR proteins determine *R* specificity (Jones and Jones, 1996; Jia et al., 2000; Dodds et al., 2001), although in some cases it is clear that additional domains are required (Luck et al., 2000). Among the LRR mutations that result in loss-of-function *rpm1* alleles, we found a slight bias toward mutation in the hydrophobic residues of the proposed LRR backbone. This fact, together with the relatively low occurrence of mutations in the LRR, leads us to believe that most positions in the LRR withstand mutation and retain *RPM1* specificity. We did not find enrichment of mutations in the third LRR, as suggested for other *R* genes (*RPS5* [Warren et al., 1998] and *RPS2* [Axtell et al., 2001]). It is possible that *RPM1* is different from these *R* genes or that our screening has a broader statistical base.

Integration of Allele Characteristics to Create Structure–Function Clusters

We sought a global understanding of the information contained in the collection of *rpm1* alleles discussed above by integrating several different parameters for each mutation. We included in this analysis the full (53) and intermediate (four) loss-of-function alleles (57 total). For each missense allele, we used four input variables in our analysis: (1) the linear position of each mutation on the RPM1 protein; (2) the PHCD between the wild-type and the mutated amino acid (Grantham, 1974); (3) the strength of the mutation [loss of function, intermediates, and intermediates that still produce cell death upon high-dose *Pst*DC3000(*avrRpm1*) inoculation]; and (4) the degree of conservation at the mutated position among 36 CC-NB-LRR proteins deduced from the Arabidopsis genome (see Methods).

As a way to integrate this information, we created clusters according to the concept of self-organizing maps (Kohonen, 1997), which are used commonly for the treatment of DNA array expression data (Maleck et al., 2000). We represented each allele within a cluster in Figure 9. On the *y* axis, we

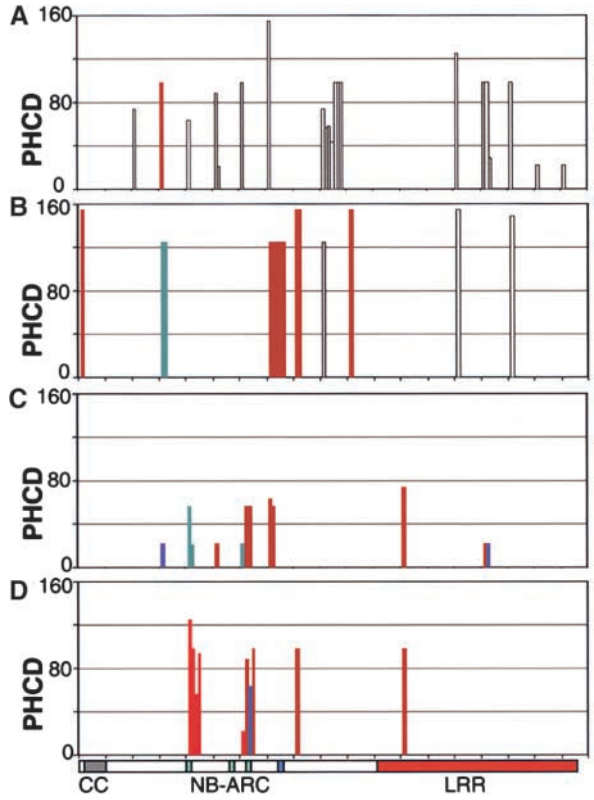

Figure 9. Clustering of the Missense Mutations Reveals Underlying Structural Correlates of RPM1 Function.

With the information presented in Table 1, we generated four clusters (shown in **[A]** to **[D]**) using the GeneCluster program. Each *rpm1* allele is represented at its linear position along the RPM1 protein (*x* axis), and the PHCD value of the substitution is shown (*y* axis). The colors represent the conservation of the wild-type amino acid at that position: red indicates 100% identical, blue indicates 100% conserved, brown indicates 75 to 99% conserved, dark cyan indicates 50 to 74% conserved, and white indicates <50% conserved. The scheme at bottom shows RPM1 with the CC motif (gray), nucleotide binding (NB-ARC; green for the canonical motifs P-loop, kinase 2, and kinase 3a [also called RNBS-B], and blue for the GLPL motif), and the LRR (red).

plotted the PHCD between the wild-type and the mutated amino acid as a measure of the severity of the mutation. We positioned each mutation along the RPM1 protein and color coded it to reflect the degree of amino acid conservation at that site (see Methods). There are several important functional inferences to be made from Figure 9. First, there is a cluster of *rpm1* alleles characterized by moderate PHCD changes in nonconserved positions (Figure 9A). Note that all four intermediate alleles are in this group and that they have the maximum PHCD within the group. These amino acids define idiosyncratic features of RPM1 that are not shared by other CC-NB-LRR proteins. We suggest that this cluster represents mutations in amino acids that might define *RPM1* specificity. A second cluster identifies positions at which a large PHCD change produces loss of RPM1 activity (Figure 9B). These sites vary in conservation from moderate to nonconserved. The third cluster is made up of very minor PHCD changes in reasonably well-conserved amino acids (Figure 9C). These positions are extremely intolerant of substitution, and it is somewhat surprising that the amino acids in this group are not preferentially the most highly conserved. This might reflect a requirement for these residues in protein stability. Finally, there is a cluster made up of moderate PHCD changes at several highly conserved amino acid positions (Figure 9D). Not surprisingly, these positions are found in or near the key domains of the NB-ARC.

Because we began by assuming that all EMS-mutable sites are equal targets in *RPM1*, our results imply that there are dozens of mutations that do not affect the *RPM1* function. We can estimate this number through the number of premature stops recovered. If all premature stops produce a loss of function (as discussed above), then the number of all possible stop codons that EMS could generate in *RPM1* (69) divided by the number of stop codons recovered and sequenced (11) provides an estimation of the total missense mutations produced. Because the total number of missense mutations that EMS could generate in *RPM1* is 715, we calculate that there were ~114 missense mutations in the screened populations. There are 57 sequenced *rpm1* missense alleles. We assume that the fraction of missense mutations in our 18 unsequenced EMS alleles is the same. Therefore, we can extrapolate that there were ~40 to 50 missense mutations in the combined EMS screens that were not identified. Although it is possible that some of these mutations could be lethal (activating alleles), we predict that there are many amino acid positions in RPM1 that are mutable and that these may be relevant for R function diversification.

METHODS

Plant Lines

Arabidopsis thaliana ecotype Columbia (*RPM1*) was used as a wild type. *rpm1-1* to *rpm1-8* alleles have been described (Grant et al., 1995). Note that *rpm1-1* has a glabrous background (Oppenheimer et al., 1991). Plants were grown in a short-day regimen as described by Ritter and Dangl (1996) and transformed as described by Clough and Bent (1998). Selection was done with the appropriate antibiotic (hygromycin) or herbicide (BASTA [McDowell et al., 1998] or chlorsulfuron [Honma et al., 1993]). The chlorsulfuron was a generous gift from DuPont. All lines discussed are single-locus T-DNA insertions.

Bacterial Strains, Inoculation, and Counting

Pseudomonas syringae pv *tomato* (*Pst*) DC3000 derivatives containing pVSP61 (empty vector, no *avr* gene), a*vrRpm1*, or *avrB* were maintained as described (Ritter and Dangl, 1996). Plant inoculations and counting of the bacteria were performed as described (Tornero and Dangl, 2001). A detailed protocol is available upon request. Where indicated, high concentrations of bacteria (OD_{600} 0.075 ~ 3.75 × 10^7 colony-forming units/mL) were infiltrated in the bottom part of the leaf with a blunt syringe for the hypersensitive response test.

Conditional Estradiol-Inducible Vector System

This inducible system requires two T-DNAs in the same plant. The first, pSGCOR1, carries a promoter from the Arabidopsis ubiquitin gene (S. Goff, unpublished results) driving the expression of a fusion protein. This fusion protein is made up of GAL4, estrogen receptor, and VP-16. The DNA fragments encoding specific domains were produced by the polymerase chain reaction (PCR) with corresponding primers to ensure in-frame translation. This T-DNA confers hygromycin resistance to the plant, and the best plant line (in terms of expression signal to noise) was called ER4. A restriction map is available upon request from S.G.

The second T-DNA, meant to express the target gene, was constructed in several steps. First, pSGCCZ1 is a pUC19 derivative that contains the inducible promoter (Guyer et al., 1998). pDG4 was created by cutting pSGCCZ1 with NotI and SalI, adding a nopaline synthase terminator sequence (from the plasmid pGPTV [Becker et al., 1992]) with borders on EcoRI and XhoI, and cloning them in pFC1 (Lonsdale et al., 1995) in the sites NotI and EcoRI. The *avrRpm1* (Dangl et al., 1992) open reading frame was amplified by PCR adding BamHI in the ATG and MluI in the stop codon. This PCR fragment was cloned in pDG4, producing pA11, and the sequence was verified. Then, pA11 was cloned into the XbaI site of pDG11, a binary vector with BASTA resistance that is a derivative from pGPTV (Becker et al., 1992), by cutting it with NheI to render pA11b. Finally, pA11b was transformed into *Agrobacterium tumefaciens* GV3101. With this strain, we transformed either wild-type plants or ER4 plants.

β-Estradiol (Sigma E 8875) was dissolved in 100% ethanol at 10 mM. These stocks are stable for at least 1 year at −20°C. Dilutions to 10 μM were made in distilled water (if the application was by spray) or 10 mM $MgCl_2$ (if the application was by infiltration of the leaf) immediately before use.

Mutagenesis and Screening

Seed were mutagenized with 0.15% ethyl methanesulfonate (EMS) for 8 hr. M2 seed were collected from 119 families of ~100 M1 plants. For T-DNA mutagenesis, the plasmid pPCVICEn4HPT (designed for activation tagging [Hayashi et al., 1992]) was modified by

replacing the *Hygromycin phosphotransferase* gene with the acetolactase synthase gene (encoding resistance to the herbicide chlorsulfuron [Honma et al., 1993]), thus creating plasmid pDG55. We pooled at least 50 T1 plants per family, producing 53 families of T-DNA mutagenized plants. M2 seeds were collected and screened independently for each family. We recovered and analyzed all putative mutants from each M2 family. We first crossed each to a11r and analyzed F1 and F2 progeny. This cross easily identifies all *rpm1* alleles, second-site mutations, and any dominant mutations. We next performed complementation analyses among those mutants that retained *RPM1* activity. Among those families that gave rise to *rpm1* alleles, we analyzed multiple siblings and never identified heterozygotes.

Genetic Analysis

The complementation groups described were obtained by standard genetic crosses and analyses of F1 and F2 progeny. These progeny were tested with β-estradiol as described above. The sequencing of *RPM1* was performed as described (Grant et al., 1995).

Clustering Analysis of the *rpm1* Alleles

To analyze the distribution of the *rpm1* mutations, we did not consider those that produce truncated proteins, internal deletions of amino acids, or frameshift mutations. We created a matrix in which, for each mutation, we considered the position of the mutation, the strength of the mutation, the physiochemical distance between the wild-type and mutant amino acid, and the level of conservation of the wild-type amino acid in RPM1 compared with all other Arabidopsis coiled-coil nucleotide binding site leucine-rich repeats (CC-NB-LRR).

The position refers to the RPM1 deduced primary structure (amino acids 1 to 926), and the strength of the mutation is obtained by assigning 0 for loss of function, 1 for intermediates, and 2 for intermediates that produce the hypersensitive response. The physiochemical distance measures the similarity between amino acids based on polarity, molecular volume, and chemical composition (Grantham, 1974). The mean distance is 100, the minimum value is 5 (isoleucine for leucine), and the maximum value is 215 (tryptophan for cysteine).

The conservation of the wild-type amino acid was obtained from the World Wide Web site http://www.tigr.org/bhaas/NBS_LRR.align.html. This represents the alignment of 36 CC-NB-LRR open reading frames (ORFs) from Arabidopsis. From this alignment, each amino acid of RPM1 was assigned a value. Those values are as follows: 5 if 100% identical, 4 if 100% conserved, 3 if 75 to 99% conserved, 2 if 50 to 74% conserved, and 1 if 0 to 50% conserved.

Once this matrix was constructed, we transformed it using Excel (Microsoft, Redmond, WA) to normalize the range of the parameters. The normalized matrix was used as input in the GeneCluster program (Tamayo et al., 1999) to obtain four clusters based on the concept of self-organizing maps (Kohonen, 1997).

Microscopy

Trypan blue staining was performed as described (Koch and Slusarenko, 1990). Microscopy slides were observed with an Eclipse E800 upright microscope (Nikon, Tokyo, Japan) equipped with a Spot charge-coupled device camera (Diagnostic Instruments, Sterling Heights, MI). The images shown were taken with a ×4 objective under normal conditions. Images were processed using the software Spot (version 2.1) and Photoshop (version 5.5; Adobe Systems, Mountain View, CA).

Accession Numbers

The GenBank accession numbers for the proteins mentioned in this article are K01486 (GAL4), XM_045966 (estrogen receptor), X03141 (VP-16), NP_187360 (RPM1), NP_194339 (RPS2), A54810 (N), and O14727 (APAF-1).

ACKNOWLEDGMENTS

We thank Syngenta for the inducible system, DuPont for the sample of chlorsulfuron, Dr. Barbara Baker for the resistance gene to chlorsulfuron, Dr. Richard Walden for the activation tagging vector, and the Molecular Pattern Recognition Group (Whitehead/Massachusetts Institute of Technology Center for Genome Research) for making available the GeneCluster software. Thanks also to members of our laboratory for their invaluable suggestions and to David Hubert and Dr. Jeff Chang for their help with the manuscript. This work was funded by National Science Foundation Grant No. IBN-9724075, renewed under the Arabidopsis 2010 project as IBN-0114795 to J.L.D., Research Experience for Undergraduates supplements to that award for R.A.C. and W.N.L., a Syngenta Collaborative Research Grant to J.L.D., and a postdoctoral fellowship from the Spanish Ministry of Science and Education to P.T.

Received September 10, 2001; accepted October 29, 2001.

REFERENCES

Aarts, N., Metz, M., Holub, E., Staskawicz, B.J., Daniels, M.J., and Parker, J.E. (1998). Different requirements for EDS1 and NDR1 by disease resistance genes define at least two R gene mediated signalling pathways in Arabidopsis. Proc. Natl. Acad. Sci. USA **95,** 10306–10311.

Aoyama, T., and Chua, N.-H. (1997). A glucocorticoid-mediated transcriptional induction system in transgenic plants. Plant J. **11,** 605–612.

Arabidopsis Genome Initiative. (2000). Analysis of the genome of the flowering plant *Arabidopsis thaliana*. Nature **408,** 796–815.

Aravind, L., Dixit, V.M., and Koonin, E.V. (1999). The domains of death: Evolution of the apoptosis machinery. Trends Biochem. **24,** 47–53.

Axtell, M.J., McNellis, T.W., Mudgett, M.B., Hsu, C.S., and Staskawicz, B.J. (2001). Mutational analysis of the *Arabidopsis RPS2* disease resistance gene and the corresponding *Pseudomonas syringae avrRpt2* avirulence gene. Mol. Plant-Microbe Interact. **14,** 181–188.

Becker, D., Kemper, E., Schell, J., and Masterson, R. (1992). New plant binary vectors with selectable markers located proximal to the left T-DNA border. Plant Mol. Biol. **20,** 1195–1197.

Bendahmane, A., Kanyuka, K., and Baulcombe, D.C. (1999). The *Rx* gene from potato controls separate virus resistance and cell death responses. Plant Cell **11,** 781–791.

Biffin, R.H. (1905). Mendel's laws of inheritance and wheat breeding. J. Agric. Sci. **1,** 4–48.

Bisgrove, S.R., Simonich, M.T., Smith, N.M., Sattler, N.M., and Innes, R.W. (1994). A disease resistance gene in Arabidopsis with specificity for two different pathogen avirulence genes. Plant Cell **6,** 927–933.

Boyes, D.C., Nam, J., and Dangl, J.L. (1998). The *Arabidopsis thaliana RPM1* disease resistance gene product is a peripheral plasma membrane protein that is degraded coincident with the hypersensitive response. Proc. Natl. Acad. Sci. USA **95,** 15849–15854.

Century, K.S., Holub, E.B., and Staskawicz, B.J. (1995). *NDR1*, a locus of *Arabidopsis thaliana* that is required for disease resistance to both a bacterial and a fungal pathogen. Proc. Natl. Acad. Sci. USA **92,** 6597–6601.

Century, K.S., Shapiro, A.D., Repetti, P.P., Dahlbeck, D., Holub, E., and Staskawicz, B.J. (1997). NDR1, a pathogen-induced component required for Arabidopsis disease resistance. Science **278,** 1963–1965.

Chen, S.-Y., Huff, S.Y., Lai, C.-C., Der, C.J., and Powers, S. (1994). Ras-15A protein shares highly similar dominant-negative biological properties with Ras-17N and forms a stable, guanine nucleotide-resistant complex with CDC25 exchange factor. Oncogene **9,** 2691–2698.

Chou, Y.L. (1969). Statistical Analysis. (New York: Holt, Rinehart, and Winston).

Clough, S.J., and Bent, A.F. (1998). Floral dip: A simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. Plant J. **16**, 735–743.

Cole, J., and Skopek, T.R. (1994). Somatic mutant frequency, mutation rates and mutational spectra in the human population in vivo. Mutat. Res. **304,** 33–105.

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

Dangl, J.L., Ritter, C., Gibbon, M.J., Mur, L.A., Wood, J.R., Goss, S., Mansfield, J., Taylor, J.D., and Vivian, A. (1992). Functional homologs of the Arabidopsis RPM1 disease resistance gene in bean and pea. Plant Cell **4**, 1359–1369.

Dinesh-Kumar, S.P., Tham, W.-H., and Baker, B.J. (2000). Structure–function analysis of the tobacco mosaic virus resistance gene *N*. Proc. Natl. Acad. Sci. USA **97,** 14789–14794.

Dodds, P.N., Lawrence, G.J., and Ellis, J.G. (2001). Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference between the *P* and *P2* rust resistance specificities in flax. Plant Cell **13**, 163–178.

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

Farnsworth, C.L., and Feig, L.A. (1991). Dominant inhibitory mutations in the Mg^{2+}-binding site of Ras^{H} prevent its activation by GTP. Mol. Cell. Biol. **11**, 4822–4829.

Feig, L.A., and Cooper, G.M. (1988). Inhibition of NIH 3T3 cell proliferation by a mutant *ras* protein with preferential affinity for GDP. Mol. Cell. Biol. **8**, 3235–3243.

Flor, H.H. (1971). Current status of the gene-for-gene concept. Annu. Rev. Phytopathol. **9,** 275–296.

Freialdnhoven, A., Scherag, B., Hollricher, K., Collinge, D.B., Thordal-Christensen, H., and Schulze-Lefert, P. (1994). Nar-1 and Nar-2, two loci required for Mla_{12}-specified race-specific resistance to powdery mildew in barley. Plant Cell **6,** 983–994.

Glazebrook, J. (2001). Genes controlling expression of defense responses in *Arabidopsis*: 2001 status. Curr. Opin. Plant Biol. **4,** 301–308.

Grant, M.R., Godiard, L., Straube, E., Ashfield, T., Lewald, J., Sattler, A., Innes, R.W., and Dangl, J.L. (1995). Structure of the *Arabidopsis RPM1* gene enabling dual specificity disease resistance. Science **269,** 843–846.

Grantham, R. (1974). Amino acid difference formula to help explain protein evolution. Science **185,** 862–864.

Guyer, D., Tuttle, A., Rouse, S., Volrath, S., Johnson, M., Potter, S., Görlach, J., Goff, S., Crossland, L., and Ward, E. (1998). Activation of latent transgenes in Arabidopsis using a hybrid transcription factor. Genetics **149,** 633–639.

Hayashi, H., Czaja, I., Lubenow, H., Schell, J., and Walden, R. (1992). Activation of a plant gene by T-DNA tagging: Auxin-independent growth in vitro. Science **258,** 1350–1352.

Heath, M.C. (2000). Hypersensitive response–related death. Plant Mol. Biol. **44,** 321–334.

Honma, M.A., Baker, B.J., and Waddell, C.S. (1993). High-frequency germinal transposition of *Ds-ALS* in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **90,** 6242–6246.

Innes, R.W., Bent, A.F., Kunkel, B.N., Bisgrove, S.R., and Staskawicz, B.J. (1993). Molecular analysis of avirulence gene *avrRpt2* and identification of a putative regulatory sequence common to all known *Pseudomonas syringae* avirulence genes. J. Bacteriol. **175,** 4859–4869.

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.

Jørgensen, J.H. (1988). Genetic analysis of barley mutants with modifications of the powdery mildew resistance gene Mla_{12}. Genome **30,** 129–132.

Kang, H.-G., Fang, Y., and Singh, K. (1999). A glucocorticoid-inducible transcription system causes severe growth defects in Arabidopsis and induces defense-related genes. Plant J. **20,** 127–133.

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

Kobe, B., and Deisenhofer, J. (1995). Proteins with leucine rich repeats. Curr. Opin. Struct. Biol. **5,** 409–416.

Koch, E., and Slusarenko, A.J. (1990). *Arabidopsis* is susceptible to infection by a downy mildew fungus. Plant Cell **2,** 437–445.

Kohonen, T. (1997). Self-Organizing Maps. (Berlin: Springer-Verlag).

Lacal, J.C. (1993). The ras Superfamily of GTPases. (Boca Raton, FL: CRC Press).

Lawrence, G., Finnegan, J., and Ellis, J. (1993). Instability of the *L6* gene for rust resistance with the presence of a linked *Ac* element. Plant J. **4,** 659–669.

Lonsdale, D.M., Moisan, L.J., and Harvey, A.J. (1995). pFC1 to pFC7: A novel family of combinatorial cloning vectors. Plant Mol. Biol. Rep. **13**, 343–345.

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.

Maleck, K., Levine, A., Eulgem, T., Morgan, A., Schmid, J., Lawton, K., Dangl, J.L., and Dietrich, R.A. (2000). The transcriptome of Arabidopsis during systemic acquired resistance. Nat. Genet. **26,** 403–410.

Malmberg, R.L. (1993). Production and analysis of plant mutants, emphasizing *Arabidopsis thaliana*. In Methods in Plant Molecular Biology and Biotechnology, B.R. Glick and J.E. Thompson, eds (Boca Raton, FL: CRC Press).

McDowell, J.M., Dhandaydham, M., Long, T.A., Aarts, M.G.M., 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.

McNellis, T.W., Mudgett, M.B., Li, K., Aoyama, T., Horvath, D., Chua, N.-H., and Staskawicz, B.J. (1998). Glucocorticoid-inducible expression of a bacterial avirulence gene in transgenic *Arabidopsis thaliana* induces hypersensitive cell death. Plant J. **14,** 247–258.

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.

Morel, J.-B., and Dangl, J.L. (1997). The hypersensitive response and the induction of cell death in plants. Cell Death Differ. **19,** 17–24.

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.

Nimchuk, Z., Rohmer, L., Chang, J.H., and Dangl, J.L. (2001). Knowing the dancer from the dance: R gene products and their interactions with other proteins from host and pathogen. Curr. Opin. Plant Biol. **4**, 288–294.

Nivard, M.J.M., Aguirrezabalaga, I., Ballering, L.A.P., Pastink, A., Sierra, L.M., and Vogel, E.M. (1999). Evaluation of the database on mutant frequencies and DNA sequence alterations of vermilion mutations induced in germ cells of Drosophila shows the importance of a neutral mutation detection system. Mutat. Res. **431,** 39–57.

Oppenheimer, D.G., Herman, P.L., Sivakumaran, S., Esch, J., and Marks, M.D. (1991). A myb gene required for leaf trichome differentiation in Arabidopsis is expressed in stipules. Cell **67,** 483–493.

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.

Ritter, C., and Dangl, J.L. (1996). Interference between two specific pathogen recognition events mediated by distinct plant disease resistance genes. Plant Cell **8,** 251–257.

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

Sega, G.A. (1984). A review of the genetic effects of ethyl methanesulfonate. Mutat. Res. **134,** 113–142.

Simonich, M.T., and Innes, R.W. (1995). A disease resistance gene in Arabidopsis with specificity for the *avrPph3* gene of *Pseudomonas syringae* pv. *phaseolicola*. Mol. Plant-Microbe Interact. **8,** 637–640.

Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., Lander, E.S., and Golub, T.R. (1999). Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. Proc. Natl. Acad. Sci. USA **96**, 2907–2912.

Tao, Y., Yuan, F., Leister, R.T., Ausubel, F.M., and Katagiri, F. (2000). Mutational analysis of the Arabidopsis nucleotide binding site–leucine-rich repeat resistance gene *RPS2*. Plant Cell **12,** 2541–2554.

Tornero, P., and Dangl, J.L. (2001). A high throughput method for quantifying growth of phytopathogenic bacteria in *Arabidopsis thaliana*. Plant J. **28,** 475–481.

Torp, J., and Jørgensen, J.H. (1986). Modification of powdery mildew resistance gene Mla_{12} by induced mutation. Can. J. Genet. Cytol. **28,** 725–731.

van der Biezen, E.A., and Jones, J.D.G. (1998). Homologies between plant resistance gene products and regulators of cell death in animals. Curr. Biol. **8,** R226–R227.

Warren, R.F., Henk, A., Mowery, P., Holub, E., and Innes, R.W. (1998). 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 **10,** 1439–1452.

Warren, R.F., Merritt, P.M., Holub, E.B., and Innes, R.W. (1999). Identification of three putative signal transduction genes involved in R gene-specified disease resistance in Arabidopsis. Genetics **152,** 401–412.

Wilhelmi, L.K., and Preuss, D. (1996). Self-sterility in Arabidopsis due to defective pollen tube guidance. Science **274,** 1535–1537.

Yu, I.-C., Parker, J., and Bent, A.F. (1998). Gene-for-gene disease resistance without the hypersensitive response in Arabidopsis dnd1 mutant. Proc. Natl. Acad. Sci. USA **95,** 7819–7824.

Zuo, J.R., and Chua, N.H. (2000). Chemical-inducible systems for regulated expression of plant genes. Curr. Opin. Biotechnol. **11**, 146–151.

Zuo, J.R., Niu, Q.W., and Chua, N.H. (2000). An estrogen receptor-based transactivator XVE mediates highly inducible gene expression in transgenic plants. Plant J. **24**, 265–273.