Published in final edited form as:

Nature. 2014 August 28; 512(7515): 436–440. doi:10.1038/nature13439.

The long–term maintenance of a resistance polymorphism through diffuse interactions

Talia L. Karasov[#1,2], **Joel M. Kniskern**[#1,†], **Liping Gao**[1], **Brody J. DeYoung**[3], **Jing Ding**[1,†], **Ullrich Dubiella**[3], **Ruben O. Lastra**[1], **Sumitha Nallu**[1], **Fabrice Roux**[4,5,6], **Roger W. Innes**[3], **Luke G. Barrett**[1,†], **Richard R. Hudson**[1], and **Joy Bergelson**[1]

[1]Department of Ecology and Evolution, University of Chicago, Chicago, Illinois 60637, USA.

[2]Committee on Genetics Genomics and Systems Biology, University of Chicago, Chicago, Illinois 60637, USA.

[3]Department of Biology, Indiana University, Bloomington, Indiana 47405, USA.

[4]INRA, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR441, F-31326 Castanet-Tolosan, France.

[5]CNRS, Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR2594, F-31326 Castanet-Tolosan, France.

[6]Laboratoire Génétique et Evolution des Populations Végétales, UMR CNRS 8198, Université des Sciences et Technologies de Lille – Lille 1, F-59655 Villeneuve d'Ascq Cedex, France.

[#] These authors contributed equally to this work.

Abstract

Plant resistance (*R*) genes are a crucial component in plant defence against pathogens[1]. Although *R* genes often fail to provide durable resistance in an agricultural context, they frequently persist as long-lived balanced polymorphisms in nature[2–4]. Standard theory explains the maintenance of such polymorphisms through a balance of the costs and benefits of resistance and virulence in a tightly coevolving host–pathogen pair[5,6]. However, many plant–pathogen interactions lack such

Author Information Reprints and permissions information is available at www.nature.com/reprints. Readers are welcome to comment on the online version of the paper.

Correspondence and requests for materials should be addressed to J.B. (jbergels@uchicago.edu)..

[†]Present addresses: Monsanto Vegetable Seeds, 37437 State Highway 16, Woodland, California 95695, USA (J.M.K.); State Key Laboratory of Plant Genetics and Germplasm Enhancement and College of Horticulture, Nanjing Agricultural University, Nanjing, China (J.D.); CSIRO Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia (L.G.B.).

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Supplementary Information is available in the online version of the paper.

Author Contributions J.B. conceived the project and organized components; J.M.K. and R.L. cloned AvrPphB2 and performed functional analyses, J.M.K. mapped and cloned *RPS5*; L.G. performed fitness trials; B.J.D. and U.D. performed effector biochemical analyses; T.L.K. performed sequence divergence analyses; J.M.K., F.R. and J.D. assayed *RPS5* frequencies in *A. thaliana*; J.M.K., S.N. and T.L.K. performed *P. syringae* fitness assays; T.L.K. cloned AvrPphB homologues; T.L.K. and R.R.H. designed the population genetics analyses and R.W.I. designed the biochemical analyses; T.L.K. performed the modelling; J.B., L.B., T.L.K. and J.M.K. were involved in the study design, the experiments and the analyses; T.L.K. and J.B. wrote the manuscript. All authors discussed the results and commented on the manuscript.

The authors declare no competing financial interests.

specificity[7]. Whether, and how, balanced polymorphisms are maintained in diffusely interacting species[8] is unknown. Here we identify anaturally interacting *R* gene and effector pair in *Arabidopsis thaliana* and its facultative plant pathogen, *Pseudomonas syringae*. The protein encoded by the *R* gene *RPS5* recognizes an AvrPphB homologue (AvrPphB2) and exhibits a balanced polymorphism that has been maintained for over 2 million years (ref. 3). Consistent with the presence of an ancient balanced polymorphism, the *R* gene confers a benefit when plants are infected with *P. syringae* carrying *avrPphB2* but also incurs a large cost in the absence of infection. *RPS5* alleles are maintained at intermediate frequencies in populations globally, suggesting ubiquitous selection for resistance. However, the presence of *P. syringae* carrying *avrPphB* is probably insufficient to explain the *RPS5* polymorphism. First, *avrPphB* homologues occur at very low frequencies in *P. syringae* populations on *A. thaliana*. Second, AvrPphB only rarely confers a virulence benefit to *P. syringae* on *A. thaliana*. Instead, we find evidence that selection for *RPS5* involves multiple non-homologous effectors and multiple pathogen species. These results and an associated model suggest that the *R* gene polymorphism in *A. thaliana* may not be maintained through a tightly coupled interaction involving a single coevolved *R* gene and effector pair. More likely, the stable polymorphism is maintained through complex and diffuse community-wide interactions.

Individuals within a population often exhibit a wide range of susceptibilities to infection and this variability has a central influence on the emergence and spread of disease[9–11]. The mechanisms that maintain resistance polymorphisms within populations have therefore been a major focus both of theoretical and empirical research[12,13]. Despite this effort, significant gaps persist in our understanding of how and why genetic variation in resistance traits is maintained. For example, current theory on the long-term maintenance of resistance polymorphisms assumes that a pathogen specializes exclusively on one host species. However, roughly half of all plant pathogens associate with multiple hosts[7], and many sparsely populated and ephemeral host species, such as *A. thaliana*, are rarely attacked by specialist pathogens. Despite the prevalence of generalist pathogens, ancient balanced polymorphisms in resistance are ubiquitous in *A. thaliana*[2]. We currently lack an understanding of how these polymorphisms are maintained.

The interaction between *A. thaliana* and the model plant pathogen *P. syringae* is an ideal system to disentangle these dynamics both because *P. syringae* is a commonly occurring pathogen in wild populations of *A. thaliana*[14] and because of the extensive knowledge of the molecular biology of this interaction[1]. *P. syringae* interacts with *A. thaliana* through the secretion of virulence-associated proteins (effectors) that downregulate plant basal defence[1]. *A. thaliana*, in turn, employs an arsenal of *R* gene products that recognize the action of specific effector proteins, inducing localized cell death (hypersensitive response) and a systemic production of chemical defences. These interactions fit the classic gene-for-gene model in requiring a close match between an effector and the corresponding *R* gene allele for a resistance response. Although *P.syringae* does not specialize on *A. thaliana*[15], multiple *R* genes in *A. thaliana* recognize specific *P. syringae* effectors and exhibit balanced resistance polymorphisms that have been maintained for millions of years[2–4]. Here we examine the dynamics of one specific gene-for-gene interaction in co-occurring *A. thaliana*

and *P. syringae* populations to identify ecological and evolutionary processes underlying the long-term maintenance of resistance polymorphisms.

We first investigated the genetic basis of recognition in the interaction between *A. thaliana* and one of its resident *P. syringae* pathogens. *P. syringae* strain PNA29.1a was isolated from a natural *A. thaliana* population[14] and shown to induce variable hypersensitive response on different *A. thaliana* accessions. To identify the effector eliciting hypersensitive response, we expressed cosmid clones of the PNA29.1a genome in a *P. syringae* strain, PstDC3000 (DC3000), which does not elicit hypersensitive response in *A.thaliana*. Agenomic fragment conferring induction of hypersensitive response to DC3000 was found to contain a homologue of the *P. syringae* effector *avrPphB* (also known as *hopAR1* and *avrPph3*). Through subcloning, we confirmed that this effector, which we name AvrPphB2, is required for recognition. To identify the corresponding *R* gene, we mapped the hypersensitive response phenotype induced by DC3000(*AvrPphB2*) on an association panel of 75 *A.thaliana* accessions[16] (Supplementary Table 1) and observed a single significant peak on chromosome1 (Fig.1a). The three most significant single nucleotide polymorphisms ($P = 1.27 \times 10^{-12}$) fell within 3.5 kilobases (kb) of *RPS5*, which encodes an R protein known to recognize AvrPphB[17]. To confirm that *RPS5* is responsible for recognition of AvrPphB2, we generated six isogenic pairs of resistant (*RPS5*$^+$) and susceptible (*RPS5*$^-$) plant lines and tested whether an AvrPphB2 induced hypersensitive response is restricted to the *RPS5*$^+$ hosts. All *RPS5*$^+$ lines expressed hypersensitive response in response to PNA29.1a after 16–24 h, whereas their *RPS5*$^-$ isolines did not. As expected, growth of DC3000(*avrPphB2*) was significantly reduced in *RPS5*$^+$ plants relative to *RPS5*$^-$ plants, even after correcting for multiple tests ($P = 1.45 \times 10^{-3}$, Wilcoxon rank-sum test) (Extended Data Fig. 1).

AvrPphB2 exhibits relatively low identity with AvrPphB at both nucleotide and protein levels (75% and 78% respectively). This divergence prompted us to ask whether AvrPphB2 triggers host resistance in a manner similar to AvrPphB by testing whether AvrPphB2 also cleaves the host protein PBS1, a virulence target of AvrPphB[18]. An immunoblot assay demonstrated cleavage of PBS1 by AvrPphB2 (Fig. 1b). Furthermore, infection of *RPS5*$^+$/*PBS1*$^-$ plants with DC3000(*avrPphB2*) failed to elicit the hypersensitive response, confirming a conserved mode of action for PBS1-dependent *RPS5* recognition of AvrPphB2. Thus, we find no evidence that sequence divergence in AvrPphB2 is associated with functional divergence in cleavage of PBS1.

RPS5 was previously shown to exhibit an ancient balanced polymorphism for its presence or absence in *A. thaliana*[3]. We confirmed this finding through a comparison of sequence divergence at the intergenic region surrounding the *RPS5* locus and here estimate that the polymorphism has been maintained in *A. thaliana* for approximately 2.6 million years (with 95% confidence intervals of 1.5–3.9 million years (refs 19, 20)). As shown previously[3], the divergences between the *Arabidopsis lyrata RPS5* locus and the *A. thaliana* presence and absence alleles are equivalent, indicating that the deletion event did not skew the inference of polymorphism age. Furthermore, we extended this work to examine polymorphisms within a broad array of *A. thaliana* populations worldwide using a PCR assay for *RPS5* presence/absence. The survey of 1,198 genetically distinct plants from 357 populations

revealed that *RPS5* was present in 51% of individuals (Fig. 2). Within the 39 populations in which we detected four or more distinct genotypes, *RPS5* was found at an average frequency of 56% (Fig. 2a), with both alleles present in more than 90% of these populations. Thus, *RPS5* polymorphism is long-lived, common in extant populations and maintained at intermediate frequency in many local populations.

This ongoing maintenance of an ancient *R* gene polymorphism within populations suggests strongly that *RPS5* confers both a fitness advantage and a fitness cost[21], probably dependent upon the presence or absence of specific pathogen effector loci. To test for a cost of resistance in the absence of infection, we measured the relative fitness of $RPS5^{+}/RPS5^{-}$ isolines in a field experiment performed in the absence of *avrPphB* homologues, including *avrPphB2*. Our experimental design included six pairs of $RPS5^{+/-}$ isolines, with replicates of each pair randomized within blocks in a field in Downer's Grove, Illinois, USA. Importantly, we used $RPS5^{+/-}$ isolines generated both in susceptible and resistant backgrounds (via Cre-lox recombination and EMS mutagenesis respectively) to allow assessment of fitness effects due to autoimmunity[22]. For the isogenic $RPS5^{+/-}$ pairs in the susceptible and resistant backgrounds respectively, the $RPS5^{+}$ genotype produced fewer seeds ($P = 0.002$, $P < 0.001$), fewer siliques ($P = 0.002$, $P < 0.001$) and less plant biomass ($P < 0.001$ both) than the $RPS5^{-}$ genotype (paired *t*-test, $n = 214$ in the susceptible background and $n = 478$ in the resistant background). The fitness cost associated with the $RPS5^{+}$ allele ranged from 5 to 10.2% (Fig. 3).

The large fitness cost associated with the presence of *RPS5* suggests that a high intensity of pathogen-mediated selection is required for the resistance polymorphism to be maintained. Consistent with this expectation, a fitness benefit of roughly 20% has been reported for plants carrying *RPS5* when infected with *P. syringae* containing avrPphB2[23]. In contrast, a survey examining the prevalence of *avrPphB* and its homologues in *P. syringae* isolates from *A. thaliana* plants in the Midwestern USA revealed that *avrPphB* homologues were restricted to 4 out of 11 populations tested, and maintained at very low frequency (6 of 183 isolates (3.27%); Fig. 2b), despite *RPS5* being found at intermediate frequencies in this region (24% of plants in US populations; Extended Data Fig. 2). *A. thaliana* was introduced relatively recently into North America, and the ancient balanced polymorphism predates this introduction. To determine whether frequencies of *avrPphB* homologues are greater in Europe where it is naturalized, we tested for the presence of AvrPphB homologues in 131 isolates from the extended *P. syringae* species complex[15] from 19 *A. thaliana* populations across France (Fig. 2c). We found putative distant homologues in three (2.29%) of these isolates. Despite the rarity of *AvrPphB* homologues, $RPS5^{+}$ is common in French populations (81% of plants in French *A. thaliana* populations possess *RPS5*). These results again suggest that a pairwise interaction between *A. thaliana* and *P. syringae* carrying homologues of *AvrPphB* may be insufficient for the maintenance of *RPS5*.

Current models to explain the maintenance of polymorphism in antagonistic species interactions typically assume tightly coupled interactions between specific host and enemy genotypes[5,6]. However, single *R* genes may recognize and interact with multiple pathogen effectors[7,24]. We hypothesized that the high frequency of *RPS5* within *A. thaliana* populations may result from selection imposed by effectors other than AvrPphB

homologues. We tested this idea by examining differential hypersensitive responses between *RPS5*$^{+/-}$ isogenic plants when infected with 42 *A. thaliana*-isolated *P. syringae* strains from the Midwestern USA. These strains were chosen to lack *avrPphB* and its homologues on the basis of dot-blot hybridizations. One strain differentially induced hypersensitive response on *RPS5*$^{+}$ plants, suggesting that *RPS5* may encode a protein that has the capacity to recognize non-homologous effectors (Supplementary Table 2). However, the low frequency of *P. syringae* isolates with any effector recognized by RPS5 in these 42 isolates (<5%) suggests additional sources of selection for resistance.

A.thaliana is host to a diverse bacterial community within its leaves[25]. A selective advantage for *RPS5* could result if multiple *A. thaliana* pathogens carry AvrPphB homologues or non-homologous effectors recognized by RPS5. In support of the former, sequence analysis suggests the potential for horizontal transfer of *avrPphB* among bacterial species. A comparison of the genomes of a *P. syringae* kiwi crop pathovar and an *A. thaliana* pathovar revealed that synonymous divergence between *avrPphB* and *avrPphB2* is substantially greater than divergence between other shared homologues (synonymous divergence = 2.28 versus an average synonymous divergence of 0.60; empirical $P = 0.003$) (Extended Data Fig. 3). The unusually ancient coalescence between *avrPphB* homologues raises the possibility that at least one AvrPphB homologue was horizontally transferred from another bacterial species. If AvrPphB homologues persist in another species, interactions between *A. thaliana* and this other species could contribute to selection for RPS5.

Spillover of pathogens from alternative neighbouring hosts has been shown to influence resistance traits in plant populations[7]. In the agricultural and disturbed ecosystems where *A. thaliana* is commonly found, *Arabidopsis* exists as a small ephemeral that is a less predictable and productive host than co-occurring perennials and crops. Other hosts can provide strong selection on the virulence of *P. syringae*; indeed, isolates of *P.syringae* collected from other hosts consistently grow in *A.thaliana* where they can demonstrate unusually high virulence[26]. Infrequent infection by highly virulent strains of *P. syringae* could impose strong selection for the ability to recognize these pathovars. It is unknown, however, if *RPS5* recognizes effectors in a wide breadth of pathovars.

To determine whether *avrPphB* and its homologues are present in strains isolated from alternative hosts, we probed a panel of 78 *P. syringae* strains isolated from crops. Successful hybridization to an *avrPphB* probe was observed for 21 out of 78 (27%) of the pathovar strains (Supplementary Table 3). We cloned 14 of these homologues, expressed them in DC3000 and tested for *RPS5*-mediated recognition. These homologues were chosen to span diverse *P. syringae* clades and thus provide a representative sample of how *A.thaliana* and *P.syringae* interact at the RPS5/avrPphB interface, at least in an agricultural context. Of course, selection imposed by crop pathovars cannot explain a polymorphism that predates the advent of agriculture. Nevertheless, these pathovars provide a window into how effectors carried by non-resident strains interact with *A. thaliana*. Approximately 70% (10 out of14) of these homologues were recognized, indicating that non-resident *P. syringae* can be subject to RPS5-mediated resistance (Extended Data Fig.4). Although the frequency of pathogenic spillover onto *A. thaliana* populations is not known, this result demonstrates that

pathovars of *P. syringae* have the potential to impose selection favouring the presence of *RPS5* in *A.thaliana* populations.

The low frequency of *avrPphB* in *P. syringae* residing within the leaves of *A. thaliana* suggests a limited fitness advantage for carrying this gene. We compared growth of DC3000 carrying an empty plasmid with the growth of DC3000(*avrPphB2*) (Extended Data Fig. 5), and found that AvrPphB2 significantly enhanced growth within an *RPS5*$^{-}$ accession (Wilcoxon rank-sum test, $P = 0.04$). The common laboratory strain DC3000 is a crop pathovar highly diverged from most *A. thaliana* isolates of *P.syringae*[27]; consequently, the virulence effect of AvrPphB2 may differ in the endemic *P. syringae* isolates. Similar experiments using each of three *P. syringae* isolates collected from *A. thaliana* populations revealed a virulence benefit in only one of the three isolates (Extended Data Fig. 6) (Wilcoxon rank-sum tests; $P = 0.01$, 0.90 and 0.50 respectively). Virulence benefits that are dependent on the genetic background help explain the relatively low frequency of *avrPphB2* within *A. thaliana* populations.

These observations are not consistent with the tight *R* gene and effector pairing thought necessary to generate stable resistance polymorphisms[4,12]. Our data instead suggest that selection favouring *RPS5* is diffuse, deriving from multiple and perhaps individually small selective agents. Similarly, the frequency of the effector in the pathogen may be only weakly coupled to the frequency of *RPS5* in *A. thaliana*. This is particularly true if *P. syringae* isolates intermittently and transiently spill over from alternative hosts to infect *A.thaliana* populations. This inconsistency between longstanding theory and our observations raises the possibility that diffuse interactions between *A. thaliana* and *P. syringae* may generate stable balanced polymorphisms.

To test this possibility, we modelled the dynamics of a resistance polymorphism under a range of parameters for a diffusely interacting host–pathogen pair. More specifically, we asked whether a balanced polymorphism could be maintained when the frequency of an *R* gene polymorphism depends upon the frequency of recognized effectors, but the frequency of effectors is uncoupled from *R* gene dynamics. To determine the conditions under which the polymorphism is stably maintained (stability is defined here as maintenance ≥ 100,000 generations), we modelled the dynamics of an *R* gene polymorphism across a range of values for the cost and the frequency of infection (Fig. 4). Next, we incorporated frequency dependence as a mechanism to couple effector and *R* gene dynamics asymmetrically without explicitly modelling effector evolution (reviewed in ref. 12). We envision a situation in which infection of susceptible plants increases resource availability for uninfected plants owing to a population-wide reduction in plant biomass or numbers. The addition of frequency dependence substantially expands the parameter space over which the *R* gene polymorphism is maintained (Fig. 4b). Across a range of infection rates (20–99%), and for relatively low costs of infection (1–30%), the balanced polymorphism can be maintained for thousands of generations (Fig. 4b and Extended Data Fig. 7). Thus, the maintenance of resistance polymorphisms does not require a tight coevolutionary pairing.

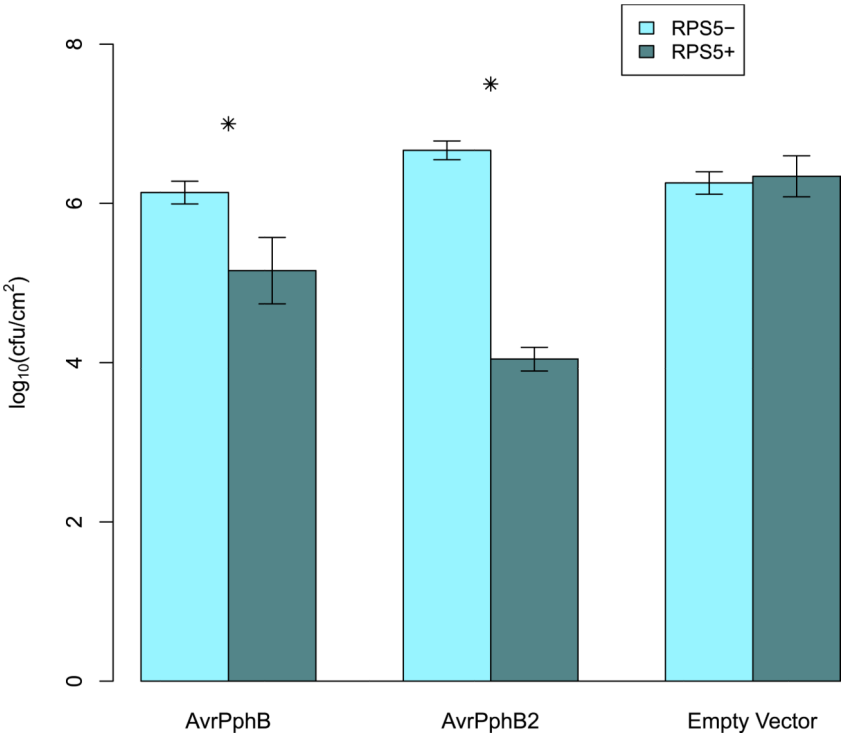
Even though it is unlikely that many *A. thaliana*–pathogen interactions are tightly coupled (but see ref.28), *A.thaliana* populations exhibit widespread balancing selection at *R* gene

loci[2]. Little is known of the ecological factors that drive *R* gene evolution in *A. thaliana*, and recent studies suggest that selection on resistance is both pathogen and host dependent. Our results point to the maintenance of one of these ancient balanced polymorphisms through complex interactions with multiple hosts sharing *P. syringae*, and through RPS5 interacting with multiple effectors distributed across multiple pathogen species. Recent studies revealing *R*-gene-mediated sexual incompatibilities and divergent selection for resistance to necrotrophs versus biotrophs suggest that numerous additional interactions may select for variation at resistance loci[22,29]. Our corresponding model demonstrates that a stable polymorphism does not require the tight coevolutionary coupling of a host and pathogen, but rather can be maintained in the face of complex interactions such as the ones we observe. Many, if not most, pathogens persist in complex ecological communities, with multiple hosts and competitors[7]. Future work on host–pathogen coevolution should elucidate whether this complexity is a general mechanism for the maintenance of variation in susceptibility and resistance.

METHODS SUMMARY

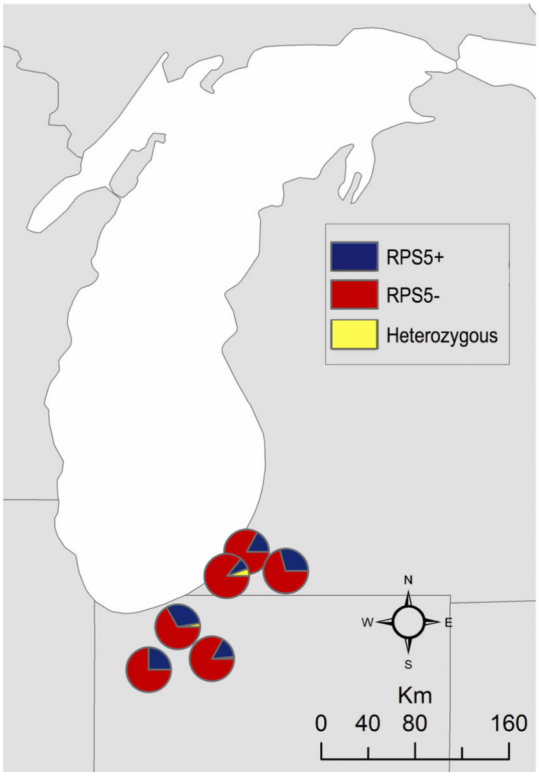
A genomic library of PNA29.1a was created using the broad-host-range vector pLAFR1, and the genomic fragment containing *avrPphB2* expressed in PstDC3000 was shown to be sufficient to induce hypersensitive response in Col-0 plants. The genetic basis of differential hypersensitive response in 75 *A. thaliana* ecotypes in response to infection with PstDC3000(*avrPphB2*) was mapped[16] using EMMAX[30]. The frequency of *RPS5* was determined by a PCR assay on genetically distinct *A. thaliana* lines. The fitness effects of the *RPS5* locus for *A. thaliana* were tested through comparisons of six paired isogenic lines. Two of the pairs were derived from Cre-lox site-specific recombination involving a fragment of the Col-0 ecotype containing the *RPS5* coding region, and the promoter and terminator regions in the Ga-0 ecotype. The remaining four pairs of isogenic lines within the Col-0 background were backcrosses of an *RPS5* mutant containing an early stop codon at amino-acid position 319 generated in the *Arabidopsis* TILLING project. To measure the fitness effects of AvrPphB2, the growth of a *P. syringae* strain transformed with an empty pME6010 plasmid was compared with strains transformed with pME6010 containing *avrPphB2*. *avrPphB* homologues were amplified and cloned into a modified version of the Gateway-compatible vector pMTN41, then expressed in DC3000. RPS5$^{+/-}$ plants were infected with DC3000 carrying the different homologues to detect differential hypersensitive response. The dynamics of the *R* gene polymorphism over time were modelled with an infinite population size and a cost of resistance of 8%. For frequency-dependent selection, the fitness of a resistant individual was modelled as increasing with the frequency of susceptible individuals.

Extended Data



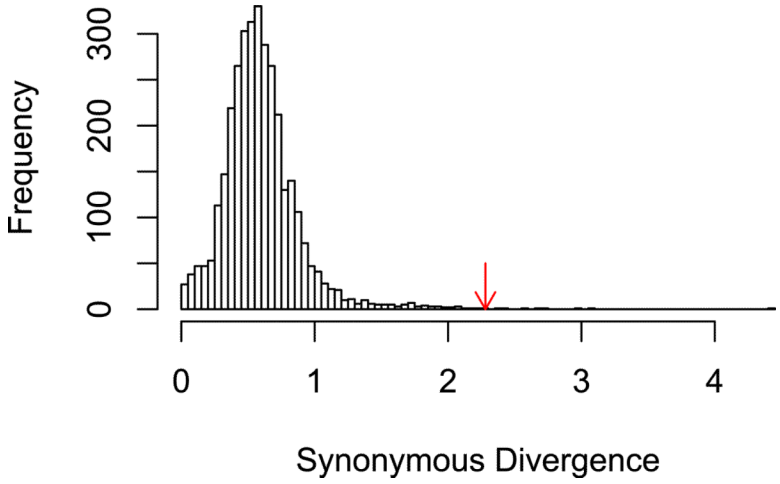
Extended Data Figure 1. Recognition of AvrPphB/2 by RPS5 reduces bacterial growth

Growth of DC3000*(avrPphB2) in planta* in Ga-0 is reduced by the presence of RPS5. In contrast, growth of DC3000 containing the empty vector pME6010 is unaffected by the presence of *RPS5*. The star denotes $P < 0.05$ in a Wilcoxon rank-sum test. Growth is measured in colony-forming units per square centimetre. Eight biological replicates were performed per genotype. Results are presented as the mean ± one s.e.m.



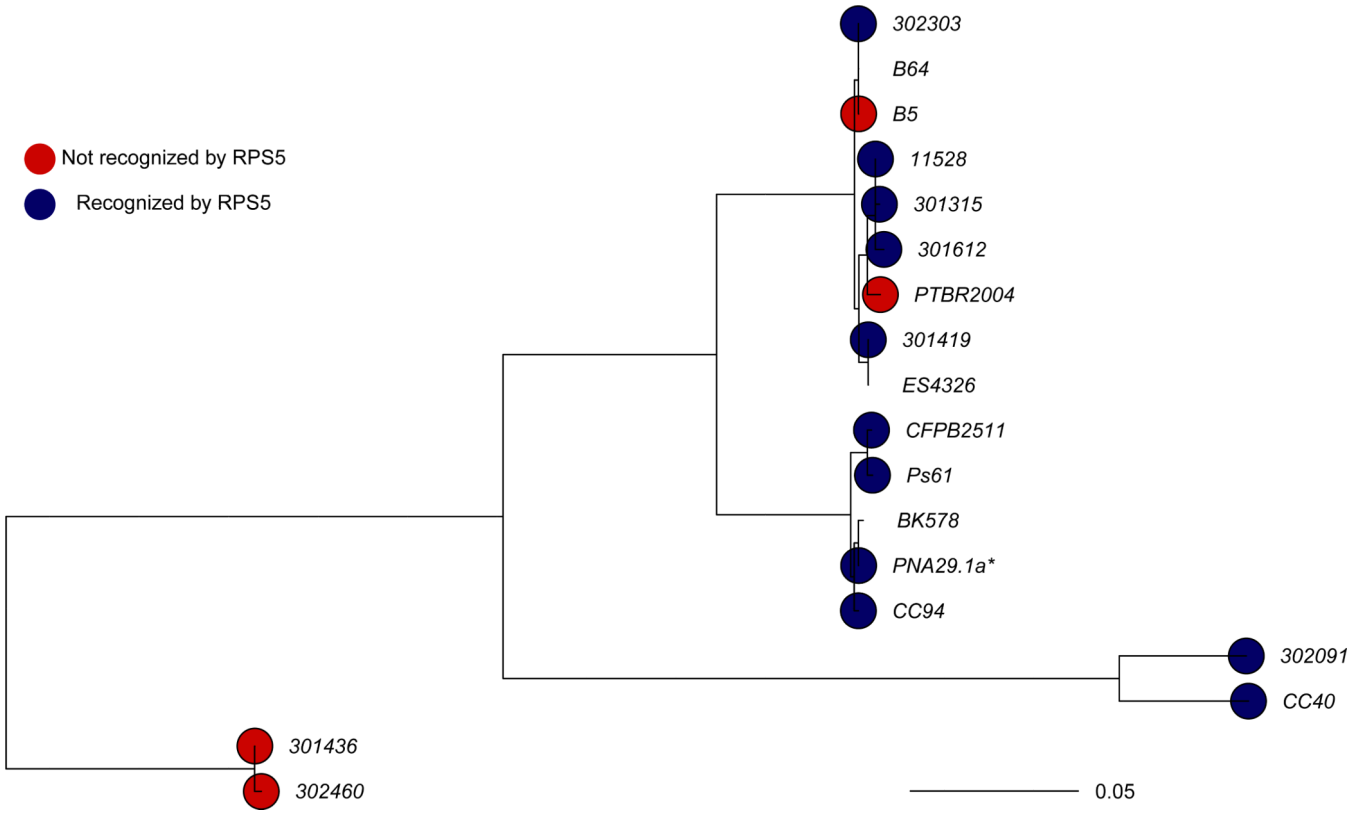
Extended Data Figure 2. Detection of RPS5 in global populations

PCR was used to test for the frequency of *RPS5* in six populations of *A. thaliana* in the Midwestern USA. The *RPS5* locus was polymorphic in all Midwestern populations. *RPS5* alleles were present at a frequency of 11–32%.



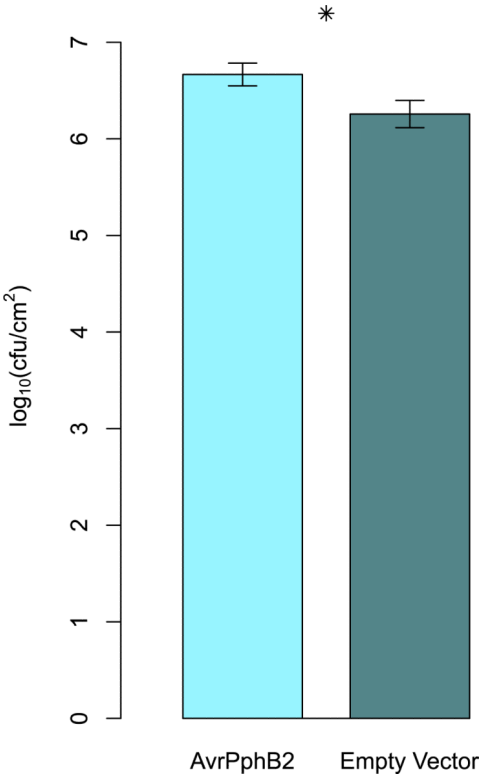
Extended Data Figure 3. Distribution of synonymous divergence in genes orthologous between *P. syringae* isolates Pan (kiwi pathovar) and PNA29.1a (*A. thaliana* pathovar)

The red arrow indicates the level of synonymous divergence between the homologues avrPphB and avrPphB2. The extreme synonymous divergence between avrPphB homologues suggests that one of the homologues has undergone horizontal gene transfer from a distantly related bacterium (empirical $P = 0.003$).



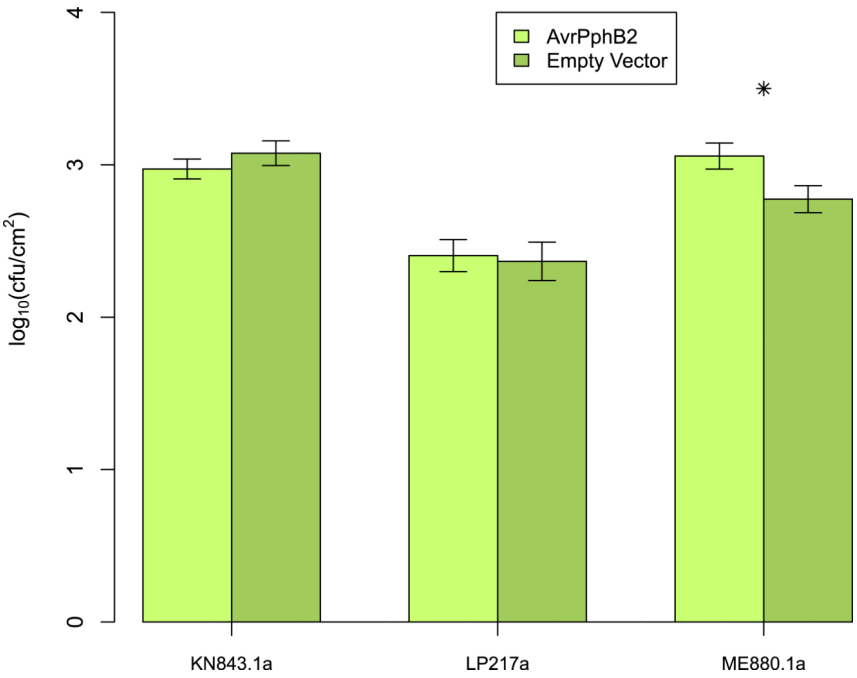
Extended Data Figure 4. AvrPphB homologues from several crop pathovars are recognized by RPS5

AvrPphB homologues found in crop pathovars were tested for the ability to elicit RPS5-mediated hypersensitive response. A maximum likelihood phylogeny of avrPphB homologues from crop pathovars and *A. thaliana* isolate PNA29.1a is presented here. The majority of homologues induced hypersensitive response. Homologues from 302460, 301436, PTBR2004 and ES4326 each encode homologues with truncated alleles. B5 encodes a full transcript. Recognition was determined by a Fisher's exact test comparison of hypersensitive response frequency upon infection of *RPS5*$^{+}$ with a homologue versus an empty vector (see Supplementary Information). The result for 302091 was marginally significant (P = 0.02, but after adjusting for multiple testing P = 0.14).



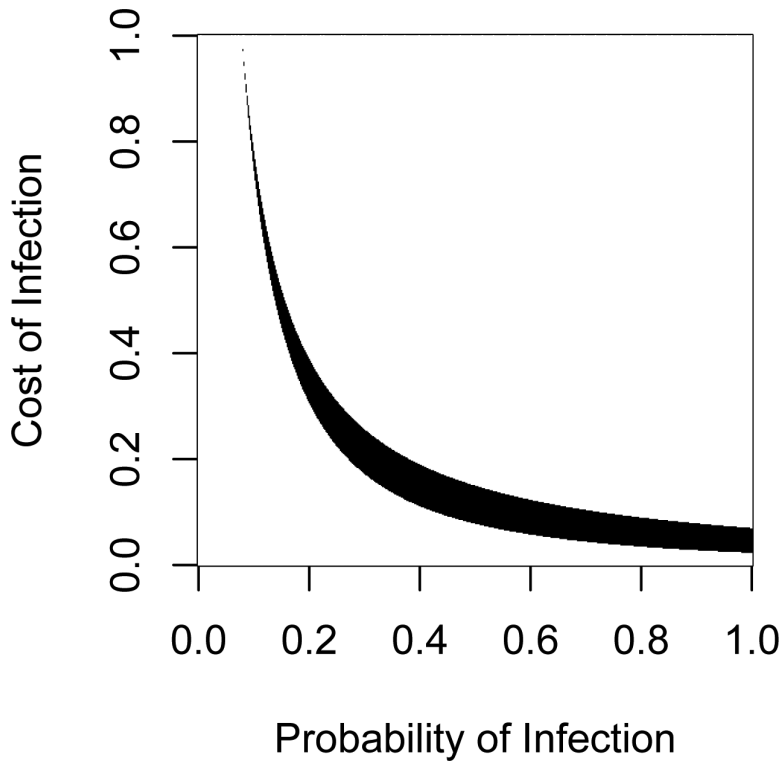
Extended Data Figure 5. AvrPphB2 enhances the proliferation of DC3000 *in planta* in the Ga-0 background

Growth of DC3000 is augmented in RPS5 plants by the presence of AvrPphB2. The star denotes $P < 0.05$ in a Wilcoxon rank-sum test. Results are presented as the mean ± one s.e.m. (calculated with seven biological replicates per genotype).



Extended Data Figure 6. The increase in virulence conferred by AvrPphB2 is genotype dependent

AvrPphB2 increases the virulence of one of three *P. syringae* isolates from *A. thaliana* populations on *RPS5*$^{-}$ Ga-0 plants. The star denotes $P < 0.0167$ (multiple-test corrected P value corresponding to $\alpha = 0.05$) in a Wilcoxon rank-sum test. Results are presented as the mean ± one s.e.m. The P values corresponding to KN843.1a, LP217a and ME880.1a are 0.401, 0.838 and 0.014 respectively (calculated with 32 biological replicates for both constructs in the KN843.1a background, 30 empty vector and 32 *avrPphB2*-containing replicates in the LP217a background and 30 empty vector, 29 *avrPphB2*-containing replicates in the ME880.1a background).



Extended Data Figure 7. Conditions for a stable polymorphism that is robust to changes in the initial frequency of the resistance allele

To determine the stability of the *R* gene polymorphism independent of the initial frequency of the *R* gene, we determined the parameters for the cost of infection and the probability of infection for which the *R* allele increases when at low frequencies but decreases at high frequencies (described in Supplementary Information). The model included frequency dependence, similar to the model used to generate Fig. 4b. The black shading signifies the conditions for which the polymorphism is robustly maintained irrespective of the starting frequency of the *R* allele.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank G. Dwyer, M. Kreitman, L. Merwin, C. Morris and M. Nordborg for discussions, D. Baltrus, D. Dahlbeck, J. Greenberg and B. Vinatzer for supplying plasmids, N. Faure for help collecting the French isolates of *P. syringae*, C. Godé, J. Higgins and A. Stathos for contributions to genotyping *RPS5* presence/absence, G. Sperone for mapping, and T. Morton and the staff of the University of Chicago greenhouse for planting and maintaining *A. thaliana*. T.L.K. was supported by a Department of Education GAANN fellowship, J.M.K. and L.G.B. were supported by postdoctoral fellowships from the Dropkin Foundation, and R.L. was supported by a National Institutes of Health (NIH) Postbaccalaureate Research Education Program award. F.R. was supported by the Laboratory of Excellence (Labex) TULIP (ANR-10-LABX-41; ANR-11-IDEX-0002-02). This work was supported by grants NIH-NIGMS R01 GM046451 to R.W.I., and grants NSF MCB0603515, NIH-NIGMS R01 GM057994 and NIH-NIGMS R01 GM083068 to J.B.

References

1. Jones JD, Dangl JL. The plant immune system. Nature. 2006; 444:323–329. [PubMed: 17108957]
2. Bakker EG, Toomajian C, Kreitman M, Bergelson J. A genome-wide survey of *R* gene polymorphisms in *Arabidopsis*. Plant Cell. 2006; 18:1803–1818. [PubMed: 16798885]
3. Tian D, Araki H, Stahl E, Bergelson J, Kreitman M. Signature of balancing selection in *Arabidopsis*. Proc. Natl Acad. Sci. USA. 2002; 99:11525–11530. [PubMed: 12172007]
4. Stahl EA, Dwyer G, Mauricio R, Kreitman M, Bergelson J. Dynamics of disease resistance polymorphism at the *Rpm1* locus of *Arabidopsis*. Nature. 1999; 400:667–671. [PubMed: 10458161]
5. Anderson RM, May RM. Coevolution of hosts and parasites. Parasitology. 1982; 85:411–426. [PubMed: 6755367]
6. Frank SA. Models of plant-pathogen coevolution. Trends Genet. 1992; 8:213–219. [PubMed: 1496557]
7. Barrett LG, Kniskern JM, Bodenhausen N, Zhang W, Bergelson J. Continua of specificity and virulence in plant host-pathogen interactions: causes and consequences. New Phytol. 2009; 183:513–529. [PubMed: 19563451]
8. Stinchcombe JR, Rausher MD. Diffuse selection on resistance to deer herbivory in the ivyleaf morning glory, *Ipomoea hederacea*. Am. Nat. 2001; 158:376–388. [PubMed: 18707334]
9. Lockett SF, et al. Mismatched human leukocyte antigen alleles protect against heterosexual HIV transmission. J. Acquir. Immune Defic. Syndr. 2001; 27:277–280. [PubMed: 11464148]
10. Paterson S, Wilson K, Pemberton JM. Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population. Proc. Natl Acad. Sci. USA. 1998; 95:3714–3719. [PubMed: 9520432]
11. Thrall PH, et al. Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. Ecol. Lett. 2012; 15:425–435. [PubMed: 22372578]
12. Brown JK, Tellier A. Plant-parasite coevolution: bridging the gap between genetics and ecology. Annu. Rev. Phytopathol. 2011; 49:345–367. [PubMed: 21513455]
13. Haldane JBS. Disease and evolution. Curr. Sci. 1992; 63:599–604.
14. Jakob K, et al. *Pseudomonas viridiflava* and *P. syringae*—natural pathogens of *Arabidopsis thaliana*. Mol. Plant Microbe Interact. 2002; 15:1195–1203. [PubMed: 12481991]
15. Morris CE, et al. Inferring the evolutionary history of the plant pathogen *Pseudomonas syringae* from its biogeography in headwaters of rivers in North America, Europe, and New Zealand. MBio. 2010; 1:e00107–10. [PubMed: 20802828]
16. Atwell S, et al. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature. 2010; 465:627–631. [PubMed: 20336072]
17. Warren RF, Merritt PM, Holub E, Innes RW. Identification of three putative signal transduction genes involved in *R* gene-specified disease resistance in *Arabidopsis*. Genetics. 1999; 152:401–412. [PubMed: 10224270]
18. Shao F, et al. Cleavage of *Arabidopsis* PBS1 by a bacterial type III effector. Science. 2003; 301:1230–1233. [PubMed: 12947197]
19. Thomson R, Pritchard JK, Shen P, Oefner PJ, Feldman MW. Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. Proc. Natl Acad. Sci. USA. 2000; 97:7360–7365. [PubMed: 10861004]
20. Hudson RR. The variance of coalescent time estimates from DNA sequences. J. Mol. Evol. 2007; 64:702–705. [PubMed: 17487522]
21. Tian D, Traw MB, Chen JQ, Kreitman M, Bergelson J. Fitness costs of *R*-gene-mediated resistance in *Arabidopsis thaliana*. Nature. 2003; 423:74–77. [PubMed: 12721627]
22. Bomblies K, et al. Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants. PLoS Biol. 2007; 5:e236. [PubMed: 17803357]
23. Gao L, Roux F, Bergelson J. Quantitative fitness effects of infection in a gene-for-gene system. New Phytol. 2009; 184:485–494. [PubMed: 19659661]
24. Grant MR, et al. Structure of the *Arabidopsis* Rpm1 gene enabling dual-specificity disease resistance. Science. 1995; 269:843–846. [PubMed: 7638602]

25. Bodenhausen N, Horton MW, Bergelson J. Bacterial communities associated with the leaves and the roots of *Arabidopsis thaliana*. PLoS ONE. 2013; 8:e56329. [PubMed: 23457551]

26. Kniskern JM, Barrett LG, Bergelson J. Maladaptation in wild populations of the generalist plant pathogen *Pseudomonas syringae*. Evolution. 2011; 65:818–830. [PubMed: 21044058]

27. Barrett LG, Bell T, Dwyer G, Bergelson J. Cheating, trade-offs and the evolution of aggressiveness in a natural pathogen population. Ecol. Lett. 2011; 14:1149–1157. [PubMed: 21951910]

28. Allen RL, et al. Host-parasite coevolutionary conflict between *Arabidopsis* and downy mildew. Science. 2004; 306:1957–1960. [PubMed: 15591208]

29. Lorang J, et al. Tricking the guard: exploiting plant defense for disease susceptibility. Science. 2012; 338:659–662. [PubMed: 23087001]

30. Kang HM, et al. Variance component model to account for sample structure in genome-wide association studies. Nature Genet. 2010; 42:348–354. [PubMed: 20208533]

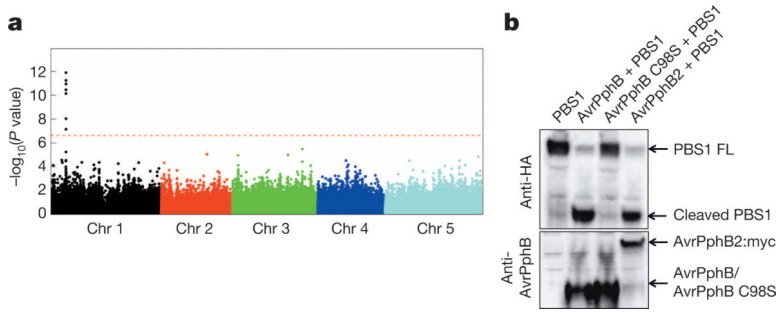


Figure 1. Identification of RPS5 and AvrPphB2 as a naturally interacting *R*-gene–effector pair in *A. thaliana* and *P. syringae* populations

a, Hypersensitive response in response to infection by DC3000(*avrPphB2*) was scored in 75 worldwide *A. thaliana* accessions that were previously genotyped at ~250,000 single nucleotide polymorphisms[16]. EMMAX[30] was used to perform genome-wide association mapping of differential hypersensitive response. The Manhattan plot illustrates the *P* values associated with each of the single nucleotide polymorphisms, and the dotted line signifies the Bonferroni correction threshold for significance ($P = 2.33 \times 10^{-27}$). The top three single nucleotide polymorphisms, each with *P* values 1.27×10^{-12}, lie within 3.5 kb of *RPS5* on chromosome (Chr) 1. **b,** AvrPphB2 cleaves PBS1. Immunoblot of PBS1 tagged with haemagglutinin (PBS1-HA) expressed alone or co-expressed with AvrPphB, AvrPphB C98S (non-active mutant) or AvrPphB2-myc in *Nicotiana benthamiana*. AvrPphB and AvrPphB2 both cleave the full-length (FL) PBS1 whereas AvrPphB C98S does not. The upper panel shows the immunoblotting results for the HA-tagged PBS1 and the bottom panel shows the results for immunoblotting for AvrPphB variants.

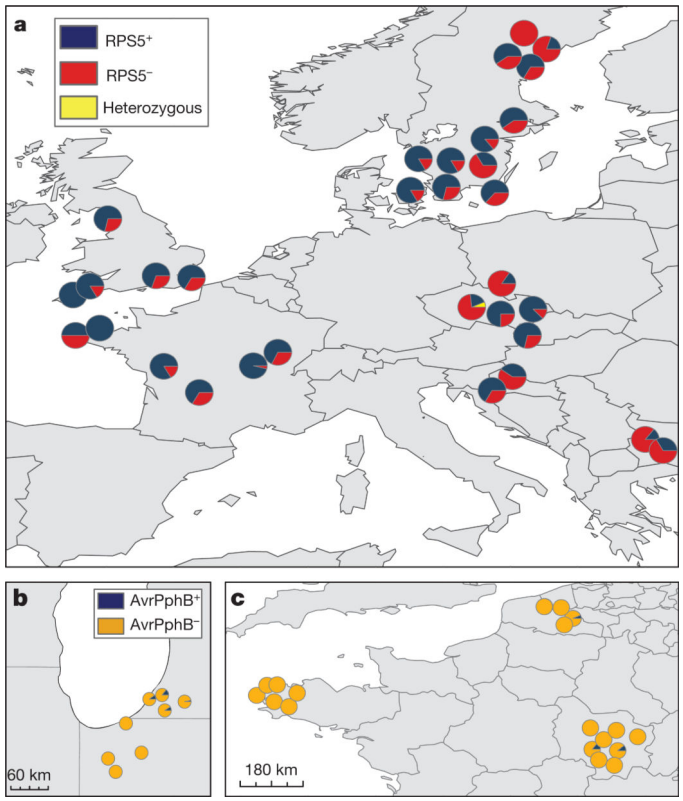


Figure 2. The distribution of *RPS5* and *avrPphB* homologues in co-occurring *A. thaliana* and *P. syringae* populations

PCR was used to test for the frequency of *RPS5* in *A. thaliana* populations across Europe (**a**) and the Midwestern USA (Extended Data Fig. 3). The *RPS5* locus was polymorphic in more than 90% (36 out of 39) of these populations (one population in eastern Asia is not illustrated). Dot-blot assays tested for the presence of AvrPphB homologues in isolates of *P. syringae* from the Midwestern USA (**b**) and from France (**c**). *avrPphB* homologues are found at 3.27% frequency in the Midwestern USA (*RPS5* frequency ~ 24%) and at 2.29% frequency in France (*RPS5* frequency ~ 81%).

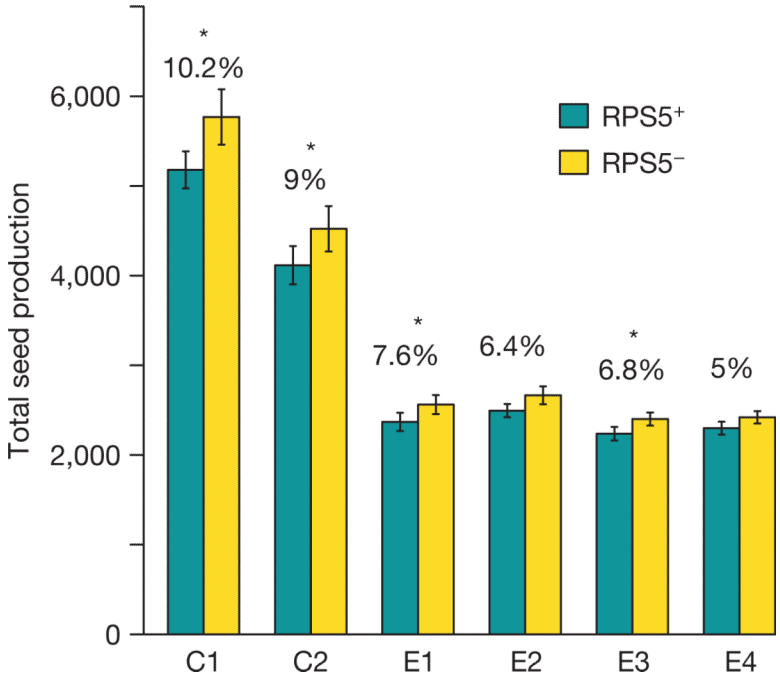


Figure 3. The cost of RPS5-mediated resistance in the absence of AvrPphB homologues

Field trials of paired isogenic lines of *A. thaliana* that differed only in the presence or absence of *RPS5* driven by its natural promoter. Lines C1 and C2 were generated by the insertion of RPS5 and its flanking regions into the susceptible Ga-0 background and subsequent excision of RPS5 using a Cre-lox system. Lines E1–E4 were generated by random point mutation in the native *RPS5* gene to generate a null allele in the resistant Col-0 background (described in Supplementary Information). In all pairings, $RPS5^{+}$ plants exhibited reduced fitness relative to $RPS5^{-}$ plants, ranging in magnitude from 5.0 to 10.2%. The percentage indicates the percentage decrease in seed production in the $RPS5^{+}$ line relative to the $RPS5^{-}$ line, and the stars denote $P < 0.05$ in a paired *t*-test (with 51 and 56 plant pairings in the two susceptible backgrounds, and 59, 52, 66 and 62 plant pairings in the resistant backgrounds). Results are presented as the mean ± one s.e.m.

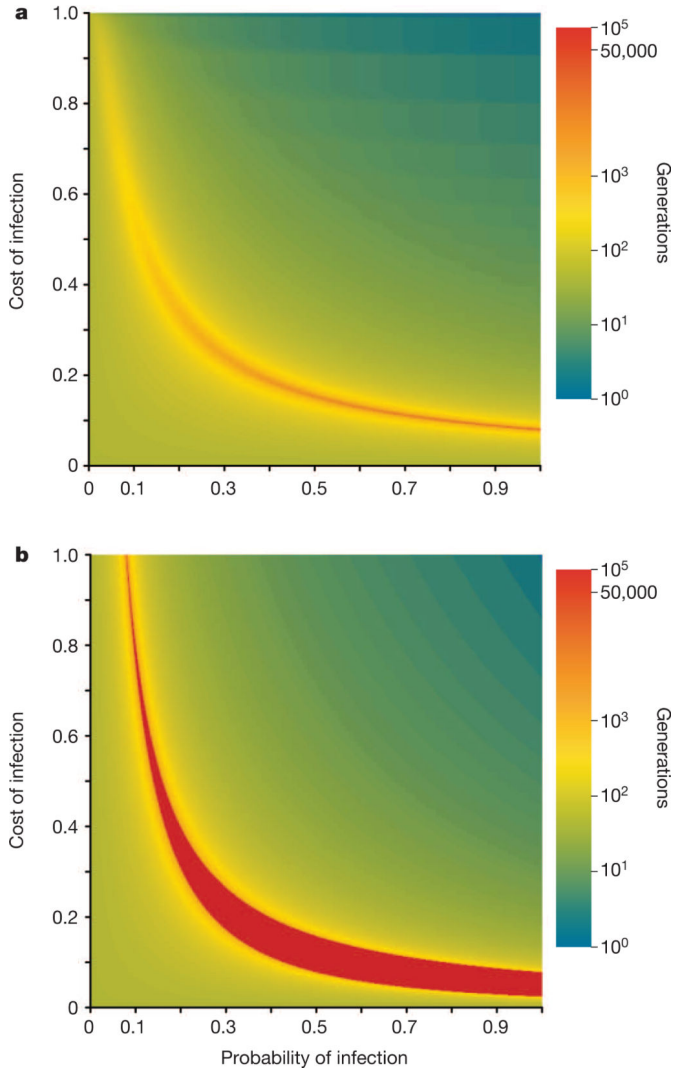


Figure 4. The maintenance of a balanced polymorphism in a diffuse interaction

To ascertain the conditions sufficient to maintain an *R* gene polymorphism stably in a diffuse interaction, we modelled the dynamics of an *R* gene polymorphism with a range of parameters for the cost of infection (*y* axis) and probability of infection with bacteria containing a recognized effector (*x* axis). With a starting frequency of 0.5 for the *R* gene and a cost of resistance of 0.08 (a value consistent with our observations in Fig. 3), we recursively simulated the frequency of the *R* gene to determine the number of generations that the *R* gene remained at a putatively detectable frequency (between 0.01 and 0.99). The simulation was run for 100,000 generations. The colour bar signifies the number of generations, up to 100,000, that the polymorphism is maintained. For **a,** the dynamics were modelled without frequency dependence. The *R* gene polymorphism is stably maintained for a narrow range of parameters. With the incorporation of negative frequency dependence into the model (described in Supplementary Information), as illustrated in **b,** the parameter space allowing a stable polymorphism expands.