(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
26 November 2009 (26.11.2009)

PCT

(10) International Publication Number
WO 2009/143133 A2

(51) International Patent Classification:
A61K 31/6615 (2006.01) *G01N 33/68* (2006.01)
A61K 38/02 (2006.01) *A61P 31/00* (2006.01)
C07K 14/44 (2006.01)

(21) International Application Number:
PCT/US2009/044489

(22) International Filing Date:
19 May 2009 (19.05.2009)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/128,080 19 May 2008 (19.05.2008) US
61/160,059 13 March 2009 (13.03.2009) US

(71) Applicant *(for all designated States except US)*: **VIRGINIA TECH INTELLECTUAL PROPERTIES, INC.** [US/US]; 2200 Kraft Drive, Suite 1050, Blacksburg, VA 24060 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **KALE, Shiv** [US/US]; 4207 Majestic Lane, Fairfax, VA 22033 (US). **TYLER, Brett** [AU/US]; 3005 Wakefield Drive, Blacksburg, VA 24060 (US). **DOU, Daolong** [CN/US]; 299 New Kent Road, Blacksburg, VA 24060 (US).

(74) Agents: **TYLER-CROSS, Ruth, E.** et al.; Whitham, Curtis, Christofferson & Cook, P.C., 11491 Sunset Hills Road, Suite 340, Reston, VA 20190 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) **Title**: COMPOSITIONS AND METHODS TO PROTECT CELLS BY BLOCKING ENTRY OF PATHOGEN PROTEINS

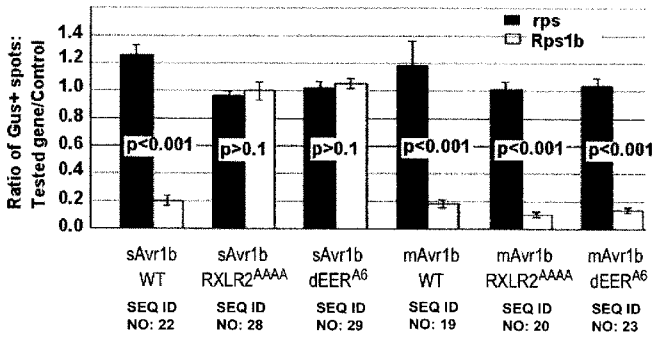


Figure 2

(57) **Abstract**: Pathogenic effector proteins which include one or more RxLR, dEER, Pexel or analogous motifs are blocked from entry into plant or animal cells by binding one or more of the motifs with a blocking compound which prevents binding of phosphoinositides or other polar lipids to the motifs which is a prerequisite for translocation of the pathogenic effector proteins into the plant or animal cell. The blocking compounds can take a variety of forms including synthetic peptides or the hydrophilic headgroups of phosphoinositides, phosphatidic acids, phospholipids, or sphingolipids. Suitable blocking compounds can be identified by assays demonstrating binding to RxLR, dEER, Pexel or analogous motifs. In addition, pathogenic effector proteins can be identified by analyzing whether they contain structural RxLR motifs using hidden markov modeling.

WO 2009/143133 A2

COMPOSITIONS AND METHODS TO PROTECT CELLS BY BLOCKING ENTRY OF PATHOGEN PROTEINS

DESCRIPTION

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention generally relates to prevention of microbial, especially oomycete or fungal, disease and, more particularly, to cellular targets for blocking entry of pathogen effector proteins into plant or animal cells. The invention also provides compositions and methods for identifying compounds that block entry of pathogen effector proteins into cells, and treatments using such compounds.

Background description

Fungi and parasites such as *Plasmodium* are eukaryotes, which are organisms that have complex internal cell structures (bacteria and viruses have simpler structures

and are excluded). Infections by parasites and fungi are especially difficult to develop drugs for because humans are also eukaryotes, so many drugs toxic to these organisms are also toxic to humans. In addition to the malaria parasite, *Plasmodium*, other eukaryotic pathogens of humans include the parasites *Schistosoma*, *Onchocerca*, *Trypanosoma*, and *Leishmania*, fungi that afflict AIDS patients such as *Candida*, *Histoplasma*, *Cryptococcus* and *Aspergillus*, and the Valley Fever fungus, *Coccidioides*, that affects healthy people in the Southwest. Fungal spores are also responsible for allergies, asthma and mold-related illnesses.

Eukaryotic pathogens of plants are also a major problem in agriculture, horticulture and forestry, and include fungi and fungal-like organisms related to marine algae called oomycetes. These diseases cause billions of dollars in losses each year. Some fungal plant pathogens include rust fungi, such as the new virulent wheat rust fungus, Ug99, that is sweeping through Africa and the middle east, and the rice blast fungus which causes major losses to the US and Asian rice crop each year. Oomycete pathogens include the late blight pathogen of potato (*Phytophthora infestans*) that causes the Irish potato famine and still causes $5 billion in losses worldwide annually, *Phytophthora ramorum* that causes Sudden Oak Death in California, and *Phytophthora sojae* that caused $1-2b damage to the US soybean crop. Worldwide transport of plants and plant products across diverse ecosystems has hastened the spread of many plant pathogens. With the increased pressure on agricultural production systems due to competing needs for food and biofuels, there is an urgent need to explore new highly efficacious strategies for biotechnology-based approaches to disease control.

Eukaryotic pathogens of both humans and plants release protein toxins called effectors that have the ability to infiltrate inside host cells, across the membrane barrier that normally surrounds the host cells. Once the effectors enter the host cell, they reprogram the cells to suppress or block the immune responses of the host and to make the host tissue more congenial for reproduction and spread of the pathogen. Therefore, drugs that could block the entry of effector proteins into host cells would

potentially suppress infection by a broad range of eukaryotic pathogens important to medicine and agriculture.

Like animals, plants have evolved defense mechanisms that afford some protection from pathogens. Constitutive defenses include structures such as the cuticle and preformed anti-microbial chemicals. Plants have also evolved an active defense response that is induced by detection of an attacking pathogen. The response includes rapid synthesis of anti-microbial chemicals and proteins, and a programmed cell death (PCD) response, called the hypersensitive response (HR). The ability of plants to detect and respond to pathogens is mediated by various receptors and signal transduction pathways that have close similarities to the innate immunity mechanisms of animals. Unfortunately, however, pathogens of both plants and animals have evolved mechanisms to avoid or suppress host defenses, thereby retaining the ability to cause many destructive diseases affecting crops and forests.

Oomycetes are fungus-like organisms many of which are pathogens. For example, most of the more than 80 species of the oomycete genus *Phytophthora* are destructive pathogens, including the potato late blight pathogen, *Phytophthora infestans*, which caused the Irish potato famine in the 18th century, the soybean root and stem rot pathogen *P. sojae*, and *Phytophthora ramorum*, the causative agent of Sudden Oak Death that is currently ravishing oak forests in California. The closely related oomycete genus *Pythium* contains more than 100 species, most of which are also pathogens. The oomycetes also include a number of commercially important and diverse downy mildew pathogens that are obligate parasites, often with narrow host ranges.

The sequencing of Expressed Sequence Tags (ESTs) and genomes from several oomycete pathogens has been completed or is under way. Draft genome sequences for the soybean pathogen *Phytophthora sojae* and for *P. ramorum* have been completed; those of *P. infestans* and the Arabidopsis downy mildew pathogen *Hyaloperonospora arabidopsidis* are nearing completion; and genome sequencing of the broad host range plant pathogens *Phytophthora capsici* and *Pythium ultimum* and the fish pathogen

Saprolegnia parasitica is also underway. In addition, substantial libraries of EST sequences are available for most of these species, as well as for *Phytophthora parasitica* and the Saprolegniomycete plant pathogens *Aphanomyces euteiches* and *Aphanomyces cochlioides*. The mining of these pathogen sequences by comparative genomics and the prediction of which proteins are secreted by the pathogens has resulted in identification of large numbers of candidate genes that potentially encode proteins involved in plant infection.

Among these are the so-called "effector proteins" or "effectors". Effectors, which are secreted by plant pathogens and have the ability to enter plant cells, have been documented for many classes of plant pathogens, including bacteria, fungi, oomycetes and nematodes. Once inside a host cell, the major function of an effector protein is to suppress the signal transduction pathways that mediate plants defense responses, and many effector proteins also suppress host programmed cell death. The activities of fungal effector proteins are known to include chitin-binding, cytotoxicity, metalloprotease activity, and protease inhibition. Pathogen effectors may also reprogram the plant cell to promote nutrition of the pathogen.

In response to pathogen attacks mediated by effectors, plants have evolved certain resistance ("R") genes that encode receptor proteins having the ability to bind and sequester, and thereby inactivate, the pathogen effector proteins. (In fact, effectors were initially discovered based on their ability to trigger responses mediated by R gene-encoded host receptors). Pathogen genes encoding effectors are referred to as avirulence (Avr) genes, because, in practice, they actually prevent infection of host plants which contain cognate receptor proteins by binding to the receptor, thereby alerting the plant to their presence, and initiating an anti-pathogen response. In contrast, genes encoding plant effectors for which cognate plant receptors do not exist are referred to as virulence genes.

Genetic mapping of oomycete Avr genes led to the cloning of the first four effector genes: Avr1b-1 from *P. sojae* (Shan, W., Cao, M., Leung, D. & Tyler, B. M. The Avr1b locus of *Phytophthora sojae* encodes an elicitor and a regulator required for avirulence on soybean plants carrying resistance gene Rps1b. Mol. Plant Microbe

Interact 17, 394-403 (2004); Avr3a from *P. infestans* (Armstrong, M. R. et al. An ancestral oomycete locus contains late blight avirulence gene Avr3a, encoding a protein that is recognized in the host cytoplasm. Proc Natl Acad Sci U S A 102, 7766-71 (2005); and ATR1 (Rehmany, A.P., Gordon, A., Rose, L.E., Allen, R.L., Armstrong, M.R., Whisson, S.C., Kamoun, S., Tyler, B.M., Birch, P.R., and Beynon, J.L. (2005). Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines. Plant Cell 17, 1839-1850), and ATR13 (Allen, R.L., Bittner-Eddy, P.D., Grenville-Briggs, L.J., Meitz, J.C., Rehmany, A.P., Rose, L.E., and Beynon, J.L. (2004). Host-parasite coevolutionary conflict between Arabidopsis and downy mildew. Science 306, 1957-1960), both from *H. arabidopsidis*. Many other effector genes have subsequently been identified and analyses have shown that all effector genes encode small secreted hydrophilic proteins that lack disulfide bonds. Significantly, effector proteins have the ability to enter plant cells unaided by any other pathogen encoded molecules. Thus, the mechanism of entry must lie in the effector proteins themselves. Sequence comparisons have led to the identification of two common motifs in the N-terminus region of effector proteins: 1) RxLR or RXLR (which stands for "arginine, any amino acid, leucine, arginine"); and 2) dEER (which stands for "aspartate which is not highly conserved, glutamate, glutamate, arginine") (Birch, P.R., Rehmany, A.P., Pritchard, L., Kamoun, S., and Beynon, J.L.2006. Trends Microbiol 14, 8-11; Tyler, 2006. Science 313, 1261-1266; Rehmany, et al. 2005. Plant Cell 17, 1839-1850). These motifs have been suspected of being responsible for the ability of effector proteins to enter plant cells. This speculation has been encouraged by the observation that a similar N-terminal "Pexel" motif (RxLxE/Q, i.e. "arginine, any amino acid, leucine, any amino acid, then aspartate or glutamate or glutamine") is required for effectors of the malaria parasite, Plasmodium, to cross the host parasitiphorous vacuolar membrane into the cytoplasm of red blood cells. In addition, experimental evidence has shown that mutations in either the RxLR or dEER motifs can alter an effector's ability to translocate into host cells.

Effectors of fungal plant pathogens have also been predicted to translocate into host cells because many plants (e.g. flax and rice) possess intracellular receptors, encoded by major resistance (R) genes, which mediate a rapid defense response when fungal effectors are present. However, prior to the present invention, fungal effectors had not been well characterized and the presence of amino acid sequence motifs that mediated entry into host cells had not been demonstrated.

In spite of previous suspicions concerning putative involvement of the RxLR and dEER motifs in effector translocation, the precise mode of and requirements for translocation of effector proteins were not known. And, in fact, it was previously not known whether fungal pathogens even possessed these or analogous motifs. This lack of knowledge had hindered the development of effective methods to combat the infection of both plant and animal host cells by oomycete, fungal and Plasmodium pathogens. Further, the lack of detailed characterization of the RxLR and dEER motifs and their flanking sequences has prevented the selection, from an enormous pool of genomic sequence data, of genes that likely encode additional effector proteins, the identification of which could lead to strategies for inhibiting their pathogenic action in cells.

SUMMARY OF THE INVENTION

The present invention provides methods to block the entry of pathogen effector proteins into host cells (e.g., "translocation"), thereby preventing host cell infection. The methods are based on the discovery that binding of phosphatidyl-inositol-3-phosphate (PI-3-P) and/or phosphatidyl-inositol- 4-phosphate (PI-4-P) and/or phosphatidic acid to effector molecules via the RxLR and dEER motif or an analogous motif is a prerequisite to translocation of the effector into a host cell, and that when binding is blocked (e.g. by inositol 1, 4-diphosphate, or any other compound which binds to one or more of RxLR and dEER motifs, such as other inositol containing phosphatidic acids, phospholipids and sphingolipids), translocation does not occur.

This strategy is also successful with fungal pathogens, since, as shown herein, the effector proteins of fungi also possess N-terminal RxLR and dEER motifs.

According to the invention, the entry of effector proteins from oomycetes, fungi, and other types of pathogens (e.g. Plasmodium) may be blocked. For example, Plasmodium effector proteins include a Pexel motif which is selectively bound as a prerequisite for translocation. Blocking of effector entry prevents the pathogen from inhibiting host cell defense mechanisms and allows the host to mount an effective response to the pathogen.

The invention also provides elucidation of the structural requirements of the RxLR and dEER motifs in oomycetes and fungi, and of the sequences which flank the motifs, leading to the ability to predict which genes in the genome of a pathogen are likely to encode effector molecules.

According to an embodiment the invention, translocation of an effector protein from a pathogen, such as a bacteria, fungus, oomycete, protozoa or nematode, into a host including animals (including humans) and plants is prevented by selectively binding a blocking compound to one or more motifs of the effector protein (e.g, RxLR, dEER, Pexel, etc.) which are bound by phosphoinositides (e.g., phosphatidyl-inositol-phosphates) or another polar lipid (e.g. phosphatidic acid) as a prerequisite for translocation. By preventing entry of the effector protein into the cell, the host cell defense mechanisms are permitted to mount an effective defense against the pathogen (it being recognized that after entry, the effector protein will compromise the host cell defense mechanisms). Thus, the invention provides a mechanism to avoid the adverse outcomes attributed to pathogenic effector proteins, and it is applicable in promoting the health and viability of both plants and animals.

Another embodiment of the invention pertains to identifying compounds which are suitable for use in protecting cells (animal and plant) from pathogenic effector proteins. In operation, an assay is used to determine whether or not a compound binds to one or more motifs of an effector protein which are bound by phosphoinositides or another polar lipid as a prerequisite for translocation. The assay may include pathogenic effector proteins which include RxLR, dEER, Pexel, or analogous motifs,

or may include protein substrates which present the RxLR, dEER, Pexel, or analogous motifs in a manner which can be bound by a candidate compound. The assay may be in the solid or liquid phase and may employ fluorescent, phosphorescent, chemiluminescent, colorimetric, or other suitable labels to indicate binding of a candidate compound to one or more motifs which are required to be bound by phosphoinositides or another polar lipid as a prerequisite for translocation.

Yet another embodiment of the invention pertains to a methodology of identifying whether an amino acid sequence of a protein in a pathogen is part of an effector protein. In this embodiment, hidden markov modeling (HMM) is used to compare flanking sequences of an RxLR sequence to determine whether the structural features for the RxLR motif are present. Through identification of effector proteins, effective strategies for preventing entry of these effector proteins into a cell (animal or plant) can be pursued.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A-E. RxLR and dEER motifs are required for Avr1b function in *P. sojae* transformants. (A) Sequences of mutations in the RxLR1, RxLR2 and dEER motifs. Bold indicates amino acids of the RxLR motifs and the alanines used to replace them in the mutations. Italics indicates the dEER motif and the alanines used to replace it in the mutant. (B) Pst I restriction analysis of PCR products amplified from Avr1b-1 transformants using primers specific for the HAM34 promoter and terminator regions. Pst I restriction profiles of Avr1b(RxLR1^{AAAA}), Avr1b(RxLR2^{AAAA}), Avr1b(RxLR1^{AAAA}, 2^{AAAA}), Avr1b(dEER^{A6}) and wild type (WT) Avr1b are distinguished from each other because the mutations introduce a Pst I site. Avr1b(dEER^{A6})-9 was confirmed by sequencing the PCR product. C, Detection of Avr1b mRNA in *P. sojae* stable transformants by RT-PCR. Upper panel shows amplification with primers internal to the Avr1b C-terminus. Lower panel shows amplification with *P. sojae* actin primers. *P. sojae* stable transformants were the same as for (B) except that an amplification reaction is also shown from RNA from a *P.*

sojae transformant containing a β-glucuronidase gene (GUS). No amplification was observed when reverse transcriptase was omitted from the reactions. (D) Distributions of HMM scores of RxLR flanking regions for all RxLR-containing secreted proteins from *P. sojae* and *P. ramorum* (non-permuted), for all secreted proteins retaining an RxLR string after sequence permutation (permuted), and for all high quality RxLR-effector candidates identified by Jiang et al (2008) (curated). The locations on the distribution of the HMM scores of the RxLR strings of known avirulence proteins and HpAvh341 are shown by the arrows. (E) Phenotype of L77-1863 (Rps1b) seedlings inoculated on the hypocotyls with transformants carrying the indicated wild type or mutant Avr1b-1 genes and photographed 4 days later.

Figure 2. RxLR and dEER functions confirmed by particle bombardment assay. Soybean leaves were bombarded using a double-barreled device that delivered Avr1b-1 DNA-bearing particles to one side of the leaf and control (empty vector) DNA to the other; both sides received GUS DNA. Ratio of blue spots in the presence of Avr1b-1 compared to the control. sAvr1b indicates a gene encoding secretory Avr1b and mAvr1b indicates one encoding mature Avr1b (lacking the secretory leader). WT indicates wild-type RxLR motif, $RxLR2^{AAAA}$ indicates the four alanine replacement of the RxLR2 motif, dEERA6 indicates the six alanine replacement of the dEER motif. Averages and standard errors are from 16 pairs of shots. p values comparing results from cultivars with Rps1b (L77-1863) or without (rps; Williams) were calculated using the Wilcoxon rank sum test.

Figure 3A-D. *P. sojae* stable transformants show that two other Avh proteins can replace the RxLR and dEER region of Avr1b. (A) Sequences of the N-termini of wild type and mutant Avr1b proteins, and of fusions with two other Avh proteins. Underlined, secretory leader; bold, RxLR motifs; italics, dEER motifs. The C-terminal sequence of Avr1b is shown in lowercase. (B) PCR analysis of DNA from *P. sojae* stable transformants. WT: pl = pHamAvr1b plasmid DNA, T17 and T20 = two transformants with wild type Avr1b-1 transgenes. HpAvh341-Avr1bCt: pl = pHamAvh341 plasmid DNA (encoding Hp Avh341-Avr1bCt), 13 and 17 = two transformants containing pHamAvh341. PsAvr4/6-Avr1bCt: pHamAvh171 plasmid

DNA (encoding Ps Avr4/6-Avr1bCt), 3 and 19 = two transformants containing pHamAvh171. mAvr1bCt: pl = pHamAvr1bCt plasmid DNA (encoding mAvr1bCt protein), 4 and 5 = two transformants containing pHamAvr1bCt. The sizes of the PCR products for Avr1b-1, pHamAvh341, pHamAvh171 and pHamAvr1bCt are 577bp, 721bp, 748bp and 385bp respectively C, Detection of Avr1b mRNA in *P. sojae* stable transformants by RT-PCR. Upper panel shows amplification with primers internal to the Avr1b C-terminus. Lower panel shows amplification with *P. sojae* actin primers. *P. sojae* stable transformants were the same as for (B) except that an amplification reaction is also shown from RNA from a *P. sojae* transformant containing a β-glucuronidase gene (GUS). pl = pHamAvr1b plasmid DNA as template. No amplification was observed when reverse transcriptase was omitted from the reactions (D) Phenotype of L77-1863 (Rps1b) seedlings inoculated on the hypocotyls with the indicated transformants carrying wild type or mutant Avr1b-1 genes and photographed 4 days later. HpAvh341-Avr1b-17, PsAvr4/6-Avr1b-3 and mAvr1bCt-5 gave similar results to HpAvh341-Avr1b-13, PsAvr4/6-Avr1b-19 and mAvr1bCt-4 (Table 1).

Figure 4A and B. Functional replacement of Avr1b host targeting signal with protein transduction motifs and Plasmodium host targeting signals. (A) Sequences of modified Avr1b proteins. PfGBP, PfHRP and Pf1615c refer to the Plasmodium Pf GBP-130, Pf HRPII and Pf PFE1615c proteins. All non-native Avr1b sequences are underlined, Avr1b RxLR2 and Plasmodium RxLXE/Q motifs are in bold, and acidic residues in the dEER region are in italics. The Avr1b secretory leader was used in all constructs. (B) Ratio of blue spots in the presence of Avr1b-1 compared to the control, assayed as described in Figure 2. Constructs are as in (A). Averages and standard errors are from 8 pairs of shots.

Figure 5. Summary of Avr1b-1 mutations and their phenotypes in P. sojae stable transformants and soybean transient expression assays. A, Avirulent; V, Virulent; NT, Not Tested; Y, significantly fewer blue (GUS-positive) tissue patches from GUS expression resulting from Avr1b-induced cell death; N, not significantly fewer blue tissue patches; P, partial reduction in blue tissue patches; SP, signal peptide.

Figures 6A-H. Binding of oomycete effector proteins to phosphoinositides. a-c, Filter-binding assays. d-f, Liposome binding assays. RxLR and dEER mutations are described in Figure 6G and H. (N)-GFP indicates a fusion of the N-terminal domain to GFP. (FL)-GST indicates a fusion of the full length effector proteins (without signal peptide) to GST. In d-f, B and F indicate liposome-bound and -free proteins respectively; M = size markers. PI-3-P = phosphatidyl inositol-3-phosphate; PI-4-P = phosphatidyl inositol-4-phosphate; PI-5-P = phosphatidyl inositol-5-phosphate; PI = phosphatidyl inositol; PA = phosphatidic acid; PS = phosphatidyl serine; PE = phosphatidyl ethanolamine; PC = phosphatidyl choline. No mutant proteins bound to PI-5-P, PI, PA, PS, PE or PC (not shown).

Figures 7A-C. Identification of host-targeting signals in fungal effectors

a, Particle bombardment cell re-entry assays of fungal effectors fused to Avr1b. N-terminal sequences of AvrL567, AvrM and AvrPi-ta (shown in b) were fused to the secretory leader (s) and C-terminal domain of Avr1b. AvrL567-Avr1b fusions lacking the secretory leader (m) or with mutations in the putative RxLR and dEER motif (rfyr-de-) were also assayed. Effector re-entry resulting in cell killing was measured by double-barreled particle bombardment in which parallel bombardments with a beta-glucuronidase (GUS) reporter gene, with and without the Avr1b fusion, were compared in the presence of resistance gene Rps1b (cultivar L77-1863) or in its absence (rps; cultivar Williams). Averages and standard errors shown are from 14-16 pairs of bombardments. P values were calculated using the Wilcoxon rank sum test. b, N-terminal sequences of effectors tested in a, with RXLR-like motifs shaded and dEER-like motifs underlined. The start of the Avr1b C-terminal domain used for all fusions is boxed. Three sequences containing potential motifs from AvrPi-ta and one from AvrL567 were inserted into Avr1b in place of the RFLR motif. Cell entry activity of each sequence is given as relative ablation. Ablation = [1 – (GUS+ spots on Rps1b)/(GUS+ spots on rps)]. Relative ablation = ablation of construct/ablation of wild-type Avr1b. In the sequence alignment, dashes indicate identical residues, periods indicate gaps in the alignment and •••• indicates Avr1b sequences.

Figures 8A-F. Binding of *P. falciparum* effector fusion proteins to phosphoinositides

a-c, Filter binding assays Mutant proteins did not bind to PI-5-P, PI, PA, PS, PE or PC (not shown) except PfGBP(N)-GFP(pexel-) which bound weakly to PS. d-f, Binding of wild-type and mutant fusion proteins to PI-3-P or PI-4-P in liposomes; B and F indicate liposome-bound and -free proteins respectively; M = size markers.

Figures 9A-D. Modulation of effector entry into root cells by phosphoinositides a, Stimulation of Avr1b(N)-GFP entry by PI-4-P and inhibition by IP2. b, Stimulation of AvrL567(N)-GFP entry by PI-4-P and inhibition by IP2. c, Entry of Arg9-GFP is not stimulated by PI-4-P nor inhibited by IP2. In each case, 1 mg/ml protein was incubated with soybean root tips for 9 hr or 12 hr then washed and photographed. Either 250 μM di-octanoyl-PI-4-P or 500 μM IP2 was preincubated with the proteins for 30 min prior to exposure to the roots. Paired light micrographs and fluorescence optical sections are from the same root tips in each case. Lighting and photographic exposure were identical for all photographs. d, Inhibition of effector-binding to liposomes by IP2. Binding of Avr1b(N)-GFP, AvrL567(N)-GFP and Arg9-GFP to liposomes containing PI-4-P was measured in the presence or absence of 300 μM inositol 1,4 diphosphate (1,4IP2). Also, binding of Arg9-GFP to liposomes containing PI-3-P was measured in the presence or absence of 300 μM inositol 1,3 diphosphate (1,3IP2).

Figures 10A-E. Effector entry into human cells and inhibition by inositol diphosphates a-d, Cells of the human lung epithelial cell line A549 were incubated with the indicated fusion proteins (1 mg/ml) for 15 hr, in the presence or absence of 430 μM inositol 1,3 diphosphate (1,3IP2), 440 μM inositol 1,4 diphosphate (1,4IP2) or 240 μM dioctanoyl-PI-4-P, then washed and photographed as described in Methods. Paired light micrographs and fluorescence optical sections are from the same cells in each case. Lighting and photographic exposure were identical for all photographs.

Figures 11A-B. Description of plasmids used in Example 1.

Figures 12A-B. Oligonucleotides used for plasmid construction. Uppercase letters indicate bases that match the initial template. Lower case letters indicate mutations or 5' extensions that do not match the initial template. Restriction sites introduced into the amplicon are underlined. A pipe (|) indicates the boundary between Avr1b-1

sequences and fused sequences (Avh, GFP or Plasmodium RXLX motif) in the fusion oligonucleotides.

Figures 13 A-F. Description of plasmids used in Example 2.

Figure 14A-F. Oligonucleotides used. Restriction sites are in bold. Mutations created by the primers are in lower case.

Figure 15. Binding of fungal effector proteins to phosphatidic acid shown in tabular form. Filter-binding assay were used to test which polar lipids were bound bind by the indicated fungal effector proteins. The N-terminus of each fungal effector (documented in the sequence list) was fused to GFP, the fusion proteins were purified from *E.coli*, and then tested for binding to the same polar lipids as documented in Figure 6. Mutant proteins contained amino acid substitutions in the motifs listed in the row above (A = alanine; S = serine). All the wild-type proteins listed bound phosphatidic acid, but none of the mutants did.

DETAILED DESCRIPTION

Since many resistance genes against oomycetes encode intracellular proteins, and since several cognate oomycete avirulence genes encode secreted proteins, it has been inferred that there must be a mechanism for translocating the avirulence proteins into the plant cells. Since the RxLR and dEER motifs were first identified during the *Phytophthora* genome sequence annotation, there has been extensive speculation that these motifs are involved in transporting avirulence proteins into host cells. Importantly, proteins in this family use the N-terminal motifs RxLR and dEER to cross the host plasma cell membrane autonomously, i.e. no other proteins are necessary to effect this translocation. Once inside the host cell, the proteins suppress host defense signaling. The importance of this effector family is underlined by the fact that plants have evolved intracellular defense receptors to detect the effectors and trigger a rapid counter-attack.

The present invention establishes that effectors of fungal plant pathogens contain functional variants of the RxLR and dEER motifs, and that the oomycete and

fungal RxLR and dEER motifs are responsible for binding of the effectors to phosphatidyl-inositol-3- phosphate (PI-3-P) and/or phosphatidyl-inositol-4-phosphate (PI-4-P) and/or phosphatidic acid. Stimulation of host cell entry by, for example, PI-4-P, and inhibition by inositol 1, 4 diphosphate suggests that the binding of polar lipids (such as phosphoinositides, phospholipids and sphingolipids) mediates cell entry of the effectors. All effectors that were tested could also enter human cells, suggesting that this mode of effector entry may be very widespread in plant and animal (including human) pathogenesis, including that which utilizes the Plasmodium Pexel ("P") motif. Identification of this broad spectrum of effectors containing functional variants of the RxLR and dEER motifs constitute a novel target class, and raises the possibility of targeted blockade of pathogen effector proteins. This knowledge can be exploited to develop new classes of antibiotic treatments to prevent a wide variety of pathogenic fungal, Plasmodial and oomycete infections in plants and animals.

The invention also identifies the sequence requirements for the function of the RxLR and dEER motifs and/or domain, and establishes that the sequences flanking the RxLR and dEER motifs are required for function. The RxLR and dEER "domain" refers to a region or regions of the primary sequence of an effector protein containing both the RxLR and dEER motifs. The sequence requirements can be defined by a hidden markov model. Mutational analysis of the RxLR motif shows that the requirement for the first and third positions are quite strict. Furthermore, reversing the order of residues 1 and 2 or of 3 and 4 also abolishes activity, indicating that the mere presence of positive charge and hydrophobicity within the motif are insufficient. The arginine at position 4 is more flexible and can be replaced by lysine or glutamine. Naturally occurring functional variants of RxLR include lysine, histidine, threonine, glycine and alanine at the fourth position.

As a result of these findings, the invention provides methods to inhibit the entry, into a host cell, of effector proteins expressed by pathogens and containing the RxLR, dEER and/or P motifs. The method is carried out by blocking the interaction, usually the binding of the motifs to a natural ligand such as a polar lipid, exemplified by phospholipids (e.g. phosphoinositides) and/or sphigolipids. Blocking may be

accomplished by any of several means, and is usually implemented by exposing the RxLR and dEER motifs to one or more molecules or molecular species which are capable of binding to or otherwise interacting with the RxLR and dEER motifs, thus preventing the polar lipid (e.g. phosphoinositide, phospholipid or sphingolipid) from binding to the motif. Herein, molecular species such as phosphoinositides, phospholipids and/or sphigolipids which, in nature, bind to one or more motifs of an effector molecule as described herein, causing the effector protein to translocate into the targeted host cell, may be referred to as "natural molecules" or "natural ligands". These designations distinguish them from the blocking molecules of the invention, which are added exogenously to cells and used to prevent binding of the natural ligands to the motifs, thereby preventing translocation of the effector protein into the cell. The blocking molecules may or may not be molecules that occur in nature, but if they are, then when used in the present invention, they are isolated or substantially purified, or chemically synthesized.

Blocking molecules of choice include but are not limited to lipid-derived molecules which bind to the motif but not in a manner that results in entry of the effector protein into the cell, e.g. molecules that are sterically related to natural ligands but which do not comprise all requisite properties for enabling translocation of the effector. In other embodiments, the blocking molecules are inositol or inositol derivatives (e.g. various phosphorylated inositols such as inositol monophosphate, various inositol diphosphates such as inositol 1, 4 diphosphate, and other similar molecules); or peptides that bind to the motif and block access to the motif by natural ligands; or peptides that bind to the motif and target the effector for protease degradation; or peptides that bind to the motif and anchor the effector to an external structure such as a cell wall or cell matrix such that the effector cannot enter the cell; or molecules that bind to the motif and cause chemical modification of the effector so that it can no longer enter cells; or other "small molecule" compounds that possess the geometric and charge requisites for binding to one or more of the motifs, thereby blocking the binding of the natural ligand that is responsible for effector translocation.

In one embodiment of the invention, the blocking molecule is a peptide, in particular a peptide with an amino acid primary sequence that is designed to include amino acid residues with charges suitable for interacting with and/or binding to the charged residues of the motif. In such a peptide, the amino acid sequence is designed so that charged atoms or groups (especially of the side chains) are spatially arranged in a manner that allows, for example, negatively charged side chains to be within bonding distance of positively charged side chains of e.g. R residues of the motif, or for aliphatic side chains of the peptide to interact with aliphatic side chains of the motif, etc. Approaches to synthetic peptide design are described, for example, by Devlin et al. (Devlin, J.J., Panganiban, L.C. and Devlin, P.E. (1990) Random peptide libraries: a source of specific protein binding molecules. Science, 249, 404-406) and Scott and Smith (Scott, J.K. and Smith, G.P. (1990) Searching for peptide ligands with an epitope library. Science, 249, 386-390). Such peptides may be designed to be stable by e.g. by avoiding the use of known protease cleavage sites in the sequence; by introducing various non-natural amino acids; or by various modifications to amino acids (e.g. amidation, sulfonation, etc.) that increase the stability of the molecule, so long a such modifications do not interfere with binding to the effector motif.

The binding or interaction of the blocking molecule(s) may be of any suitable type, and will depend on the nature of the blocking molecule. For example, the binding may be covalent and hence essentially irreversible. Thus, in some embodiments of the invention, the blocking molecule is one that, upon contact with one or more chemically reactive functional groups of the motif, forms a covalent bond with the one or more functional groups, or with functional groups of adjacent residues of the protein, in a manner that blocks access to the motif by e.g. phospholipids and/or sphingolipids that are natural ligands of the motif, and which, upon binding to the motif, permit translocation of the effector into the host cell that the pathogen is trying to infect. Usually, however, the binding is non-covalent and comprises, for example, electrostatic and/or charge interactions, hydrophobic interactions, van der Waals interactions, etc. For a blocking molecule to be effective, a Kd better than ten-fold less than the concentration of the competing natural ligand in the region of the host

membrane is preferred (a lower Kd indicates tighter binding). In this manner, binding of the natural ligand is prevented or at least attenuated or slowed so as to render the natural ligands ineffective in enabling the effector molecule to enter the targeted host cell, and infection of the host by the pathogen which manufactured the effector molecule is prevented, or attenuated or slowed. Those of skill in the art will recognize that much benefit can accrue from a treatment that inhibits a pathogen, even if inhibition is not absolute, but merely attenuates or slows the symptoms of infection. "Inhibition" or "prevention of infection" as used herein is intended to encompass all such degrees of inhibition, blocking, etc. In addition, in the case of plants, the overall biomass of harvested plants or plant products, and therefore the usefulness of the crop, may be advantageously increased even if some plants remain affected by the pathogen after treatment as described herein.

"Motifs" to which natural ligand binding is blocked include RxLR motifs, dEER motifs, and the Pexel motif. Those of skill in the art will recognize that many effector proteins contain both an RxLR motif and a dEER motif. According to the invention, in one embodiment, the binding of a natural ligand such as a phospholipid or sphingolipid to either one or the other, or both, of the RxLR and dEER motifs is blocked, and blocking occurs in a manner that prevents the effector protein that bears the motif(s) from entering the host cell. In another embodiment, the binding of a natural ligand to a Pexel motif is blocked, and blocking occurs in a manner that prevents the effector protein that bears the motif from entering a host cell. In some embodiments, the blocking molecule binds to or interacts directly with residues of one or both of the RxLR and dEER motifs, or the Pexel motif. However, this need not always be the case. In some embodiment, the blocking molecule binds to or interacts with adjacent residues. "Adjacent residues" may, but need not necessarily be, adjacent in primary sequence to the motif. They may also be in proximity due to the secondary or tertiary structure of the effector molecule. For example, in one embodiment, the effector protein comprises an RXLR motif followed by at least one aspartate or one glutamate residue within a 60 amino acid carboxy terminal flanking sequence. In other words, within the effector protein, the sequence which is attached directly to the

carboxy terminus of the RxLR motif (which, in primary sequence, follows immediately after the carboxyl terminal R of the motif) contains at least one aspartate residue and/or at least one glutamate residue within the first 60 amino acids of the sequence.

In addition, the RxLR motifs that are targeted for blocking by the methods of the invention include but are not limited to those which comprise at least one of a two or three amino acid sequence selected from the group consisting of: arginine, any amino acid, leucine; histidine, any amino acid, leucine; lysine, any amino acid, leucine; arginine, any amino acid, isoleucine; histidine, any amino acid, isoleucine; lysine, any amino acid, isoleucine; arginine, any amino acid, methionine; histidine, any amino acid, methionine; lysine, any amino acid, methionine; arginine, any amino acid, tyrosine; histidine, any amino acid, tyrosine; lysine, any amino acid, tyrosine; arginine, any amino acid, phenylalanine; histidine, any amino acid, phenylalanine; lysine, any amino acid, phenylalanine; arginine, any amino acid, tryptophan; histidine, any amino acid, tryptophan; lysine, any amino acid, tryptophan; arginine, any amino acid, valine; histidine, any amino acid, valine; lysine, any amino acid, valine; arginine, leucine; histidine, leucine; lysine, leucine; arginine, isoleucine; histidine, isoleucine; lysine, isoleucine; arginine, methionine; histidine, methionine; lysine, methionine; arginine, tyrosine; histidine, tyrosine; lysine, tyrosine; arginine, phenylalanine; histidine, phenylalanine; lysine, phenylalanine; arginine, tryptophan; histidine, tryptophan; lysine, tryptophan; arginine, valine; histidine, valine; and lysine, valine. These particular sequence may be represented by standard conventions using the single letter abbreviation for the amino acid and an "x" for the variable residue, e.g. as RXL for "arginine, any amino acid, leucine".

The host cells that are protected from effector protein invasion include many species of plant and animal cells, including human cells. Examples of plant cells that can benefit from the practice of the invention include but are not limited to: wheat, maize, rice, sorghum, barley, oats, millet, soybean, common bean (e.g. *Phaseolus* species), green pea (*Pisum* species), cowpea, chickpea, alfalfa, clover, tomato, potato, tobacco, pepper, egg plant, grape, strawberry, raspberry, cranberry, blueberry,

blackberry, hops, walnut, apple, peach, plum, pistachio, apricot, almond, pear, avocado, cacao, coffee, tea, pineapple, passionfruit, coconut, date and oil palm, citrus, safflower, carrot, sesame, common bean, banana, citrus (e.g. orange, lemon, grapefruit), papaya, macadamia, guava, pomegranate, pecan, *Brassica* species (canola, cabbage, cauliflower, mustard etc), cucurbits (pumpkin, cantaloupe, squash, zucchini, melons etc), cotton, sugar cane, sugar beets, sunflower, lettuce, onion, garlic, ornamental cut flowers, grasses used in lawns, athletic fields, golf courses and pastures (e.g. *Festuca*, *Lolium*, *Zoysia*, *Agrostis*, *Cynodon*, *Dactylis*, *Phleum*, *Phalaris*, *Poa*, *Bromua* and *Agropyron* species), etc.

Examples of animal cells that may benefit from the practice of the invention include but are not limited to: humans, cattle, sheep, pigs, goats, horses, cats, dogs, chickens, turkeys, bees, salmon, trout, bass, catfish, shellfish, crayfish, lobsters, shrimp, crabs, etc.

Many types of invasive pathogens may be targeted and their effector proteins prevented from entering host cells by the methods of the invention. Examples of such pathogens include but are not limited to: any *Phytophthora* species, e.g. *Phytophthora infestans*, *Phytophthora sojae*, *Phytophthora ramorum*, *Phytophthora parasitica*, *Phytophthora capsici*, *Phytophthora nicotianae*, *Phytophthora cinnamomi*, *Phytophthora cryptogea*, *Phytophthora drechsleri*, *Phytophthora cactorum*, *Phytophthora cambivora*, *Phytophthora citrophthora*, *Phytophthora citricola*, *Phytophthora megasperma*, *Phytophthora palmivora*, *Phytophthora megakarya*, *Phytophthora boehmeriae*, *Phytophthora kernoviae*, *Phytophthora erythroseptica*, *Phytophthora fragariae*, *Phytophthora heveae*, *Phytophthora lateralis*, *Phytophthora syringae*; any *Pythium* species, e.g. *Pythium ultimum*, *Pythium aphanidermatum*, *Pythium irregulare*, *Pythium graminicola*, *Pythium arrhenomanes*, *Pythium insidiosum*; any downy mildew species; any *Peronospora* species, e.g. *Peronospora tabacina*, *Peronospora destructor*, *Peronospora sparsa*, *Peronospora viciae*; any *Bremia* species, e.g. *Bremia lactucae*; any *Plasmopora* species, e.g. *Plasmopora viticola*, *Plasmopara halstedii*; any *Pseudoperonospora* species, e.g. *Pseudoperonospora cubensis*, *Pseudoperonospora humuli*; any *Sclerospora* species

e.g. *Sclerospora graminicola*; any *Peronosclerospora* species, e.g. *Peronosclerospora philippinesis*, *Peronosclerospora sorghi*, *Peronosclerospora sacchari*; any *Sclerophthora* species, e.g. *Sclerophthora rayssiae*, *Sclerophthora macrospora*; any *Albugo* species, e.g. *Albugo candida*; any *Aphanomyces* species, e.g. *Aphanomyces cochlioides*, *Aphanomyces euteiches*, *Aphanomyces invadans*; any *Saprolegnia* species, e.g. *Saprolegnia parasitica*; any *Achlya* species; any rust fungi; any smut fungi; any bunt fungi; any powdery mildew fungi; any *Puccinia* species, *Puccinia striiformis*, *Puccinia graminis*, *Puccinia triticina* (syn. *Puccinia recondita*), *Puccinia sorghi*, *Puccinia schedonnardii*, *Puccinia cacabata*; any *Phakopsora* species, e.g. *Phakopsora pachyrhizi*, *Phakopsora gossypii*; any *Phoma* species, e.g. *Phoma glycinicola*; any *Ascochyta* species, e.g. *Ascochyta gossypii*; any *Cryphonectria* species, e.g. *Cryphonectria parasitica*; any *Magnaporthe* species, e.g. *Magnaporthe oryzae*; any *Gaeumannomyces* species, e.g. *Gaeumannomyces graminis*; any *Synchytrium* species, e.g. Synchytrium endobioticum; any *Ustilago* species, e.g. *Ustilago maydis*, *Ustilago tritici*, *Ustilaginoidea virens*; any *Tilletia* species, e.g. *Tilletia indica*, *Tilletia caries*, *Tilletia foetida*, *Tilletia barclayana*; any *Erysiphe* species, e.g. *Erysiphe necator* (formerly *Uncinula necator*); any *Blumeria* species, e.g. *Blumeria graminis*; *Podosphaera oxyacanthae*; any *Alternaria* species, e.g. *Alternaria alternata*; any *Botrytis* species, e.g. *Botrytis cinerea*; any *Diaporthe* species, e.g. *Diaporthe phaseolorum*; any *Fusarium* species, e.g. *Fusarium graminearum*, *Fusarium oxysporum*, *Fusarium moniliforme*, *Fusarium solani*; any *Leptosphaeria* species, e.g. *Leptosphaeria maculans*, *Leptosphaeria maydis*; any *Macrophomina* species, e.g. *Macrophomina phaseolina*; any *Monilinia* species, e.g. *Monilinia fructicola*; any *Mycosphaerella* species, e.g. *Mycosphaerella graminicola*, *Mycosphaerella fijiensis*, *Mycosphaerella tassiana*, *Mycosphaerella zeae-maydis*; any *Phialophora species*, e.g. *Phialophora gregata*; any *Phymatotrichopsis species*, e.g. *Phymatotrichopsis omnivora*; any *Taphrina* species, e.g. *Taphrina deformans*; any *Aspergillus* species, e.g. *Aspergillus flavus*, *Aspergillus parasiticus*, *Aspergillus fumigatus*; any *Verticillium* species, e.g. *Verticillium dahliae*, *Verticillium albo-atrum*, *Rhizoctonia solani*, *Ophiostoma ulmi* (syn. *Ceratocystis ulmi*), *Ophiostoma novo-ulmi*; any

Septoria species, e.g. *Septoria avenae*; any *Pyrenophora* species, e.g. *Pyrenophora tritici-repentis*; any *Colletotrichum* species, e.g. *Colletotrichum graminicola*; any *Sclerotinia* species, e.g. *Sclerotinia sclerotiorum*; any *Sclerotium* species, e.g *Sclerotium rolfsii*; any *Thielaviopsis* species, e.g *Thielaviopsis basicola*; any *Coccidioides* species, e.g. *Coccidioides immitus*; any *Paracoccidioides* species, e.g. *Paracoccidioides braziliensis*; any *Pneumocystis* species, e.g. *Pneumocystis carinii*; any *Histoplasma* species, e.g. *Histoplasma capsulatum*; any *Cryptococcus* species, e.g. *Cryptococcus neoformans*; any *Candida* species, e.g. *Candida albicans*; any apicomplexan parasite species such as: any *Plasmodium* species, e.g. *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae*; any *Babesia* species, e.g. *Babesia bovis*, *Babesia bigemina*; any *Cryptosporidium* species, e.g. *Cryptosporidium parvum*; any *Toxoplasma* species, e.g. *Toxoplasma gondii*; any *Trypanosomatid* species such as: any *Trypanosoma* species, e.g. *Trypanosoma brucei*, *Trypanosoma cruzi*, *Trypanosoma congolense*, *Trypanosoma vivax*; any *Leishmania* species, e.g. *Leismania donovani*. Any amebozoan parasites; any *Entamoeba* species, e.g. *Entamoeba histolytica*; any *Mastigamoeba* species; any *Schistosoma* species; any *Onchocerca* species; any *Giardia* species; any microsporidial species; any *Enterocytozoon* species; any *Encephalitozoon* species, e.g. *Encephalitozoon cuniculi*, etc.

As a result of the practice of the methods of the invention, entry of a pathogen effector protein into the host cell is prevented, inhibited, slowed, or otherwise decreased or lessened. As a result, the host cell can mount a robust or normal immune response to the pathogen, and infection of the host organism is averted (prevented), or the degree of infection (i.e. the deleterious symptoms that typically accompany the presence of an infection by the pathogen) are eliminated or decreased. Thus, some aspects of the invention also include methods of preventing or attenuating the symptoms of infection usually caused in a host organism by a pathogen which employs effector proteins comprising the RxLR and dEER motif to enter host cells. In some embodiments, the methods are used to prevent infection and/or symptoms of infection. In other embodiments, infection may have already started but the methods of the

invention can be used to curtail the spread of the infection to other organisms, or to lessen the symptoms in an organism that is already afflicted. This is the case, in particular, with *Plasmodium* infections, where the methods of the invention are especially useful in preventing the subsequent rounds of parasite multiplication after initial infection, or with other pathogens that multiply logarithmically. In addition, the invention provides methods of maintaining a host cell's ability to mount an immune response to a pathogen, the pathogen being one that produces effector proteins that comprise one or more of the motifs described herein, and the method involving blocking the effector protein from entering the host cell by preventing the binding of its natural ligand.

Administration of the blocking molecules of the invention will depend on several factors, including the nature of the molecule and the host. Generally, the blocking molecule will be in a composition or formulation suitable for administration. If the host organism is a plant, application is generally in the form of a foliar spray or watering solution of e.g. an aqueous or oil solution that includes the blocking molecule in a concentration sufficient to block effector molecules of pathogens which are likely to attack the plant. For administration to an animal, which may be a human, any suitable composition, many of which are known in the art, may be employed, e.g. various pills, powders, liquids, injectable formulations, etc. Likewise, any suitable means may be used, including but not limited to by injection (e.g. subcutaneous or intramuscular), inhalation, orally, intranasally, by ingestion of a food product containing the protein, etc. In addition, the compositions may include one or more than one blocking molecule. For example, a preparation for application to plants may include molecules that block the effector proteins of one or of several different types of pathogen. In addition, the blocking molecules may be administered to plants in conjunction with other beneficial substances, such as fertilizers, various pesticides, growth factors, etc. The same is true for administration to animals, where on or more than one type of blocking molecule may be administered, and may be administered in conjunction with other beneficial substances such as chemotherapeutic agents that also have activity against the pathogen.

In another aspect, the invention provides elucidation of the structural requirements of the RXLR and dEER motifs in oomycetes and fungi, and of the sequences which flank the motifs, thereby allowing, for example, the design of molecules to bind the motif. As a result, the invention also provides methods to predict which genes in the genome of a pathogen are likely to encode effector molecules. This is significant because, as demonstrated herein, the mere presence of a sequence conforming to the RXLR and dEER motif in a protein is not sufficient to insure that the protein is an effector, i.e. that the protein is able to traverse the cell wall and enter the cell upon binding to phospholipids or sphingolipids. Functional RxLR and dEER motifs (or a functional RxLR and dEER domain) have additional requirements, particularly in the flanking sequences, as described herein. Briefly, the invention describes a non-random distribution of amino acid residues in the regions flanking the RXLR motif, represented by a position-weight matrix. One method of predicting whether or not a gene encodes a true effector protein involves the use of a hidden markov model (HMM) based on the position-weight matrix.

The development of this aspect of the invention was based in the observation of differences in activity between two putative RXLR motifs, RXLR1 and RXLR2, as described in detail in Example 1 below. The differences suggested that surrounding sequences are also important to the activity of an RXLR motif. To define the differences in the surrounding sequences a hidden markov model (HMM) was created using the 10 amino acid residues to the left and right of the RXLR motifs of all of known *P. sojae* and *P. ramorum* Avh genes. Using this HMM, the sequences surrounding RXLR2 (which was an authentic, functional motif) had a high score of 18.5, representing an excellent match to the consensus flanking sequence, and very unlikely to have been found in a random sequence. In contrast, sequences surrounding RXLR1 (a non-functional motif) had a low, non-significant score of 0.0. Using the same HMM, the sequences surrounding the RXLR motif of *P. infestans* Avr3a scored 10.9. Using a similar HMM derived from *Hyaloperonospora parasitica* Avh genes, the sequences surrounding the RXLR motifs of the *H. parasitica* Atr1 and Atr13 proteins had scores of 9.8 and 6.3 respectively. These and other comparisons described in

Example 1, showed that HMM scores of zero, such as that of Avr1b RXLR1, are characteristic of RXLR strings found at random, i.e. such sequences are not likely to represent true RXLR motifs. On the other hand, HMM scores over 5.0 are characteristic of non-random occurrences of RXLR strings, and the proteins in which such non-random strings occur are likely to be authentic functional RXLR sequences. In other words, such sequences are likely, upon binding a phospholipid or sphingolipid, to promote or allow the translocation, into a host cell, of effector (avirulence) proteins in which they are located or of which they form a part. HMM scores between 0 and 5 are equivocal and cannot be assigned to either category of protein (random RXLR strings vs authentic RXLR motifs).

For further discussions of the use of a hidden markov model, see United States patent 6,128,587 to Sjolander (October 3, 2000), the complete contents of which are hereby incorporated by reference.

The invention also provides a method for screening compounds to identify those that inhibit binding of phosphoinositides, phospholipids or sphingolipids (or other natural lipid ligands) to one or both of an RXLR and a dEER motif in an effector protein, and which thus can inhibit translocation of effector proteins that have these motifs into cells. The method involves exposing a candidate or putative blocking compound to one or both of the RXLR and dEER motifs under conditions suitable for binding of either the natural phospholipid/sphingolipid ligands or phosphoinositides to the motifs (phosphoinositides are known to be capable of such binding). The screening is evaluated in that if the candidate compound is able to bind to one or both of the RXLR and dEER motifs, then the compound is selected as a compound that will inhibit binding of the natural ligands to the motifs in an effector protein, and prevent entry of the effector protein into a cell. This is especially the case if the blocking compound is able to competitively bind to the motif in the presence of a natural ligand.

The invention is further illustrated by the following Examples, which should not be interpreted as limiting the invention in any way.

EXAMPLES

EXAMPLE 1. RXLR-mediated entry of *Phytophthora sojae* effector Avr1b (SEQ ID NO: 2) into soybean cells does not require pathogen encoded machinery

Effector proteins secreted by oomycete and fungal pathogens have been inferred to enter host cells, where they interact with host resistance gene products. Using the effector protein Avr1b of *Phytophthora sojae*, an oomycete pathogen of soybean, we show that a pair of sequence motifs, RXLR and dEER, plus surrounding sequences, (SEQ ID NO: 46) are both necessary and sufficient to deliver the protein into plant cells. Particle bombardment experiments demonstrate that these motifs function in the absence of the pathogen, indicating that no additional pathogen encoded machinery is required for effector protein entry into host cells. Furthermore, fusion of the Avr1b RXLR and dEER domain to green fluorescent protein (GFP) allows GFP to enter soybean root cells autonomously. The conclusion that RXLR and dEER serve to transduce oomycete effectors into host cells indicates that the more than 370 RXLR and dEER containing proteins encoded in the genome sequence of *P. sojae* are candidate effectors. We further show that the RXLR and dEER motifs can be replaced by the closely related erythrocyte targeting signals found in effector proteins of Plasmodium, the protozoan that causes malaria in humans. Mutational analysis of the RXLR motif shows that the required residues are very similar in the motifs of Plasmodium and Phytophthora. Thus the machinery of the hosts (soybean and human) targeted by the effectors may be very ancient.

RXLR2 and dEER motifs of Avr1b are required for its avirulence function in transgenic P. sojae lines

To test the function of the RXLR and dEER motifs of Avr1b, transgenic *P. sojae* strains that expressed either wild-type or mutant Avr1b-1 genes were created. Wild-type Avr1b contains two RXLR motifs, RXLR1 and RXLR2 (Figure 1A). Mutations in either or both of the RXLR motifs (SEQ ID NO: 3,4,5), in addition to a mutation in the dEER motif (Figure 1A) (SEQ ID NO: 6), were created. The Avr1b-1 gene constructs were fused to a strong constitutive promoter, HAM34 (Judelson, H., Tyler, B.M., and Michelmore, R.W.1991. Mol. Plant-Microbe Interact. 4, 602-607.),

and introduced into a strain, P7076, that expresses a variant Avr1b protein that does not confer avirulence against Rps1b-containing soybeans (Shan, W., Cao, M., Leung, D., and Tyler, B.M. 2004. Mol. Plant-Microbe Interact. 17, 394-403). Two independent transformants (T17 and T20) expressing wild type Avr1b-1 (Figure 1B, C) lost the ability to infect soybean plants carrying Rps1b, but were unaffected in their ability to infect plants lacking Rps1b (Figure 1E, Table 1). This demonstrated that they had acquired avirulence against Rps1b as a result of a functional Avr1b gene product. This result was confirmed using two different pairs of isolines of soybean that differed only in the presence of Rps1b, namely Williams (no Rps gene) with L77-1863 (Rps1b; Williams background) and HARO(1-7)1 (No Rps; Harosoy background) with HARO13 (Rps1b; Harosoy background) (Figure 1E; Table 1).

In contrast, in five independent transformants expressing the RXLR2^{AAAA} (SEQ ID NO 4)mutant, there was no gain of avirulence against Rps1b cultivars, despite the presence of abundant mRNA from the transgene (Figure 1E, Table 1). Thus, the RXLR2 motif is necessary for Avr1b activity when the protein is delivered by the pathogen. Since the RXLR1 motif was intact in the RXLR2^{AAAA} mutant, the motif appeared to be non-functional. Consistent with this inference, the RXLR1^{AAAA} mutation (SEQ ID NO 3) did not abolish avirulence in three independent transformants (Figure 1E, Table 1). As expected, avirulence was lost in the RXLR1^{AAAA}, RXLR2^{AAAA} double mutants (Figure 1E, Table 1) (SEQ ID NO: 5). A mutation in the dEER motif (SEQ ID NO: 6) also abolished avirulence (in two independent transformants) indicating that this motif is also required for the function of the protein (Figure 1E, Table 1).

Table 1. Molecular characterization and avirulence testing of *P. sojae* stable transformants

Strains	Transgene (PCR)[a]	Expression (RT-PCR)[b]	Transformant Validation[c]	Avirulence (Surviving seedlings)[d]	p[e]

				Rps1[b]	rps	
P7076 (Gus)						
GUS	no	no		1/23	1/22	0.77
P7076 (*sAvr1b* WT)						
T17	yes	yes	Pst I	20/30	0/21	3.9 E-07
T20	yes	yes	Sequence	25/44	2/28	1.2 E-05
P7076 (*sAvr1b* RXLR1^{AAAA})						
RXLR1-2	yes	yes	Pst I	32/36	3/20	5.0 E-08
RXLR1-3	yes	yes	Pst I	31/54	3/21	6.4 E-04
RXLR1-5	yes	yes	Pst I	37/57	7/23	5.1 E-03
P7076 (*sAvr1b* RXLR2^{AAAA})						
RXLR2-18	yes	yes	Pst I	6/44	3/17	0.48
RXLR2-20	yes	yes	Pst I	4/46	3/19	0.33
P7076 (*sAvr1b* RXLR1^{AAAA}; RXLR2^{AAAA})						
RXLR1+2-4	yes	yes	Pst	5/31	2/16	0.55
RXLR1+2-6	yes	yes	Pst I	4/43	1/21	0.47
P7076 (*sAvr1b* dEER^{A6})						
dEER-9	yes	yes	Sequence	4/59	2/23	0.79
dEER-14	yes	yes	Pst I	4/40	3/15	0.28
Ps *Avr4/6-Avr1bCt*						
4/6-1b-3	yes	yes	size	15/24	3/11	0.057

4/6-1b-19	yes	yes	size	14/26	3/16	0.025
Hp *341-Avr1bCt*						
341-1b-13	yes	yes	size	16/19	0/11	6.6 E-06
341-1b-17	yes	yes	size	23/24	0/7	3.0 E-06
mAvr1bCt						
mAvr1bCt-4	yes	yes	size	3/20	3/16	0.77
mAvr1bCt-5	yes	yes	size	2/21	1/15	0.63

[a] The presence of transgenes was verified by PCR as described in the Materials and Methods. + = transgene present; - = transgene not detected.

[b] Transgene expression was determined by qualitative RT-PCR (RT-PCR) and by quantitative RT-PCR (q-PCR) as described in the Materials and Methods. Yes = transgene transcripts present; no = transgene transcripts not detected; ND = not determined.

[c] The presence of the relevant mutation in the transforming plasmid was verified by sequencing in every case. The presence of the correct mutation within the transgenes of each transformed strain was verified after PCR amplification of the Avr1b-1 transgene by Pst I digestion or by sequencing in the case of the mutants (e.g. Figure 1) or by size in the case of the Avh gene fusions and N-terminal deletion (e.g. Figure 4).

[d] The avirulence of each transgenic strain was tested by inoculation of seedlings containing Rps1b (L77-1863) or no rps gene (Williams), as described in the Materials and Methods. The number of surviving seedlings/total inoculated seedlings is shown, summed from all replicates

[e] Fisher's exact test (one tailed) was used to compare the frequency of seedling survival between *rps* and *Rps1b* plants. A significant p value (0.05) indicates that the transformant's phenotype is avirulent.

The difference in activity between RXLR1 (SEQ ID NO: 3) and RXLR2 (SEQ ID NO: 4) suggests that surrounding sequences are important to the activity of an RXLR motif. To define the differences in the surrounding sequences a hidden markov model (HMM) was created using the 10 amino acid residues to the left and right of the RXLR motifs of all of the *P. sojae* and *P. ramorum* Avh genes (Tyler et al., 2006; Jiang et al., 2008). Using this HMM, the sequences surrounding RXLR2 (SEQ ID NO: 5) had a high score of 18.5, representing an excellent match to the consensus flanking sequence, very unlikely to have been found in a random sequence. In contrast, sequences surrounding RXLR1 (SEQ ID NO: 3) had a low, non-significant score of 0.0. Using the same HMM, the sequences surrounding the RXLR motif of *P. infestans* Avr3a scored 10.9. Using a similar HMM derived from *Hyaloperonospora parasitica* Avh genes, the sequences surrounding the RXLR motifs of the *H. parasitica* Atr1 and Atr13 proteins had scores of 9.8 and 6.3 respectively.

To establish the significance of these HMM scores, the *Phytophthora* HMM was used to score the RXLR motifs of 1240 RXLR-containing sequences identified from a pool of all putative secreted *P. sojae* and *P. ramorum* proteins by Jiang et al. (2008). As a control, 639 RXLR-containing sequences were scored found after permuting the sequences of all the putative secreted *P. sojae* and *P. ramorum* proteins (Jiang et al., 2008). As shown in Figure 1D, the RXLR strings of 698 (56%) of the 1240 real proteins had an HMM score of zero, while the RXLR strings of 595 (93%) of the permuted proteins had a zero score, and only 13 (1.8%) scored above 5.0. In contrast, of 765 proteins that Jiang et al. (2008) identified as high quality candidate effectors, only 18% had an HMM score of zero, and 543 (72%) had a score over 5.0. From this comparison we conclude that HMM scores of zero, such as that of Avr1b RXLR1 (SEQ ID NO: 3), are characteristic of RXLR strings found at random, while scores over 5.0 are characteristic of non-random occurrences of RXLR strings and of the RXLR strings of functional avirulence proteins. HMM scores between 0 and 5 are equivocal. The curated Avh genes with a score of zero may represent pseudogenes as many of them were identified principally by C-terminal sequence similarity.

The interaction between Avr1b and the Rps1b gene product occurs within host cells and does not require the RXLR and dEER motifs

To confirm that the site of interaction of Avr1b with the Rps1b gene product is within the plant cell, particle bombardment was used to introduce DNA encoding Avr1b proteins lacking a secretory leader into soybean cells together with DNA encoding ß-glucuronidase (GUS). This assay measures the functional interaction of the Avr1b protein with the intracellular product of the soybean Rps1b gene; when the two proteins interact, programmed cell death is triggered in the transformed cells ablating the development of tissue patches expressing GUS. Since the Avr1b protein lacks its normal secretory leader, the protein should be synthesized in the plant cytoplasm. To facilitate the comparison of test and control bombardments, a novel double-barreled attachment for the Bio-Rad Gene Gun was utilized (Dou et al., 2008). The gun shooting two different DNA samples side-by-side into a leaf in the same shot, which greatly improves the reproducibility of the results (Dou et al., 2008). Figure 2 shows that delivery of DNA encoding leader-less Avr1b protein (SEQ ID NO: 19) into soybean cells significantly reduced the number of blue GUS-positive patches when the Rps1b gene was present, but not when Rps1b was absent (Figure 2A). This is consistent with a cytoplasmic location for the Avr1b-Rps1b interaction. When RXLR2 or dEER motifs were replaced by four or six alanine residues, respectively [Figure 2A, mAvr1b(RXLR2^{AAAA}) (SEQ ID NO: 20) and mAvr1b(dEER^{A6})(SEQ ID NO: 23)], the interaction of the cytoplasmic, leader-less Avr1b with Rps1b was unaffected (Figure 2A), indicating that the RXLR2 and dEER motifs were not required for the interaction.

RXLR-mediated transit into soybean cells does not require the pathogen

To test whether RXLR function requires the presence of the pathogen, the bombardment assay was used to determine the effect of the RXLR2^{AAAA} mutation on secreted Avr1b protein (SEQ ID NO: 18). When soybean cells were bombarded with DNA encoding wild type Avr1b (SEQ ID NO: 22), including its normal secretory leader, a reduction in GUS-positive blue spots was observed comparable to that observed for the non-secreted protein [Figure 2, sAvr1b(WT)]. However when the RXLR2^{AAAA} (SEQ ID NO: 28)or dEERA6 (SEQ ID NO: 29) mutations were present in

the bombarded DNA, there was no reduction in the number of blue spots [Figure 2, sAvr1b(RXLR2^{AAAA}) (SEQ ID NO: 11) and sAvr1b(dEER^{A6}) (SEQ ID NO: 13)]. From these results it may be inferred first that the secretory leader is functional in soybean and targets Avr1b protein to the outside of the cell; second, that the RXLR2 (but not RXLR1) and dEER motifs are required for Avr1b protein to re-enter the cell, which confirms the conclusion from the *P. sojae* transformation experiments. Importantly, the results also show that RXLR-dEER-mediated entry does not require the presence of the pathogen.

To support the inference that the secretory leader of Avr1b was correctly exporting the protein from the plant cells in the bombardment assay, a gene encoding *Aequorea coerulescens* green fluorescent protein (acGFP; "GFP" herein) fused either to the Avr1b leader (SEQ ID NO: 27) or to full-length Avr1b (SEQ ID NO: 25) was constructed. These fusions enabled tracking of the proteins and checking their stability. To aid in visualization onion bulb epidermal cells were used rather than soybean cells. GFP was exported from the cells and accumulated in the apoplast when the secretory leader was attached to GFP but accumulated in the cytoplasm and nucleus when the leader was not attached (not shown). When full length Avr1b was fused to GFP (SEQ ID NO: 25), the proteins also accumulated in the apoplast if a mutation was present in either the RXLR2 motif (SEQ ID No: 28) or in the dEER motif (SEQ ID No: 29). This observation confirmed that the protein encoded by these mutants was stable and correctly targeted outside of the cells. When cells expressing Avr1b-GFP fusion proteins with RXLR mutations were plasmolyzed by treatment with 0.8M mannitol for 15 min, the GFP was associated with the cell wall and not with the plasma cell membrane (not shown). Furthermore, GFP protein could be seen diffusing into the apoplast between pairs of neighboring cells not shown). Similar observations were made when cells expressing secreted GFP or Avr1b-GFP fusion proteins with a dEER mutation were plasmolyzed (not shown). If the RXLR2 and dEER motifs were intact however, the sAvr1b-GFP protein fusion accumulated in the cytoplasm and nucleus of the cells, similar to the mAvr1b-GFP (SEQ ID NO: 24) fusion lacking the leader. When cells expressing sAvr1b-GFP fusion proteins were plasmolyzed by

treatment with 0.8M mannitol for 15 min, the GFP could be observed to have either fully or partially returned to the inside of the cells. These results supported our conclusion that the RXLR2 and dEER motifs act together to enable Avr1b protein to re-enter the plant cells.

The Avr1b RXLR and dEER motifs are sufficient to target GFP to soybean cells

The RXLR and dEER region of Avr1b was fused to GFP (SEQ ID NO: 46), and the fusion protein was synthesized in *E. coli* and partially purified. Root tips of soybean seedlings were incubated with the isolated fusion protein for 12 hours, washed for four hours in water, then observed under light and UV microscopy to localize the GFP. GFP accumulated inside many of the root cells, whereas buffer alone did not produce any fluorescence. The optical sections produced by the confocal microscope revealed that the protein penetrated approximately 10 cell layers deep during the 12 hour incubation. The characteristic accumulation of GFP in the nuclei of the treated cells is comparable to the pattern observed when GFP is expressed *in planta*, and verifies that the GFP is located inside the cells. The nuclear localization of the protein also indicates that the cells are alive. If mutations were present in the RXLR or dEER motifs of the fusion protein, GFP did not accumulate inside the soybean root cells. When the RXLR and dEER region was replaced by the artificial protein transduction motif Arg9 (SEQ ID NO: 112), GFP once again entered the soybean root cells and accumulated in the nuclei.

Avr1b RXLR and dEER motifs can be replaced by RXLR and dEER-containing protein sequences encoded by bioinformatically identified Avh genes

To determine if the RXLR and dEER motifs of bioinformatically identified Avh genes could functionally replace the RXLR2 and dEER motifs of *Avr1b-1*, full length Avh genes from *P. sojae* and *H. parasitica* were fused to an Avr1b-1 N-terminal deletion mutant lacking the RXLR and dEER motifs. The fusion genes were then introduced into *P. sojae* and the transformants were tested for avirulence on Rps1b-containing soybean cultivars. Both Avh genes, *P. sojae* Avh171 (since identified as Avr4/6; Dou et al., 2008) (SEQ ID NO: 8) and *H. parasitica* Avh341 (SEQ ID NO: 7), could replace the requirement for the RXLR2 and dEER motifs as judged by the

avirulence of the transformants on Rps1b-containing cultivars, whereas transformants containing only the C-terminus of Avr1b fused to an initiator methionine remained virulent (Figure 3 and Table 1). This result indicates that the RXLR and dEER motifs form a distinct transferable functional domain of Avr1b and other Avh proteins. The HMM scores of the RXLR-dEER motifs of Ps Avr4/6 and Hp Avh341 are both well within the functional range (6.9 and 14.2, respectively).

The Avr1b host targeting signal can be functionally replaced by autonomous protein transduction motifs

Protein transduction domains (PTDs) capable of autonomously carrying proteins across plasma cell membranes have been described and characterized in the HIV-1 Tat protein. Arginine-rich peptides such as Arg9 can also carry out this function. To compare RXLR and dEER mediated effector delivery with the function of PTDs, the RXLR2 motif of Avr1b was replaced with the TAT PTD (SEQ ID NO: 42) or with Arg9 (Figure 4A) (SEQ ID NO: 41). The resultant proteins were treated using the particle bombardment assay, and both PTDs could functionally replace the RXLR2 motif of Avr1b, restoring the avirulence reaction of Avr1b with Rps1b (Figure 4B). Furthermore, when the version of secreted Avr1b that contained the Arg9 sequence in place of the RXLR2 motif was fused to GFP (SEQ ID NO: 30), the fusion protein accumulated in the cytoplasm and the nucleus of bombarded onion bulb cells rather than the apoplast, confirming that Arg9 could functionally replace RXLR2 (not shown). Similar results were obtained when the version of secreted Avr1b that contained the TAT PTD was fused in place of the RXLR2 motif to GFP (SEQ ID NO: 31). Finally, when Arg9 was fused to GFP (SEQ ID NO: 112), the isolated proteins could enter soybean root cells directly (not shown).

The Avr1b host targeting signal is interchangeable with host targeting signals from Plasmodium effectors

To test if the erythrocyte targeting signals of Plasmodium effector proteins could functionally replace the RXLR and dEER region of Avr1b, the residues of Avr1b from the end of the secretory leader to the end of the dEER motif were replaced with the mature N-termini of three different *Plasmodium* effector proteins that are

targeted to the erythrocyte cytoplasm, namely PfGBP-130 (SEQ ID NO: 121), PfHRPII (SEQ ID NO: 123) and PfPFE1615c (SEQ ID NO: 125) (Bhattacharjee, S., Hiller, N.L., Liolios, K., Win, J., Kanneganti, T.D., Young, C., Kamoun, S., and Haldar, K. 2006. PLoS Pathog 2, e50.). The entire 37-41 amino acid region of each *Plasmodium* effector required for transduction was used (Figure 4A). As shown in Figure 4B, all three *Plasmodium* host targeting domains could functionally replace the Avr1b N-terminus in targeting Avr1b to the soybean cytoplasm, assuming that they do not simply interfere with secretion.

Functional characterization of the RXLR motif

To experimentally characterize the sequence requirements of the RXLR motif, a series of mutations were introduced into the motif in a version of the Avr1b-1 gene that retained the secretory leader, and assayed the mutants using the bombardment assay (Table 2). Mutations which targeted the arginine at position 1 or the leucine at position 3 has the strongest effect on the ability of Avr1b to ablate GUS-positive tissue patches. Replacement of R1 with lysine reduced function significantly (33% ablation compared to 78%; $p < 0.001$) while glutamine replacement completely abolished it. Replacement of L3 with alanine or even the relatively conservative valine also completely abolished function. Replacement of the arginine at position 4 with a glutamine slightly but significantly reduced function (58% ablation compared to 72%; $p < 0.001$). Reversing the order within the first and second two pairs of positively-charged and hydrophobic residues (RFLR -> FRLR; RFLR -> RFRL) completely abolished avirulence activity, indicating that positions of R1 and L3 were critical, not just their presence.

Table 2. Function of RXLR2 mutants of Avr1b assayed by particle bombardment

SEQ ID NO	RXLR2 sequence[a]	Ratio of GUS-positive spots[b]		ablation[c]	p value[d]	Activity
		rps	Rps1b			

22	RFLR	1.26 ± 0.07	0.28 ± 0.03	0.78	< 0.001	Yes
20	AAAA	0.93 ± 0.04	0.96 ± 0.05	0	>0.1	No
38	KFLR	1.04 ± 0.04	0.70 ± 0.04	0.33e	< 0.001	Partial
37	QFLR	0.95 ± 0.03	0.99 ± 0.03	0	>0.1	No
31	FRLR	1.00 ± 0.04	0.98 ± 0.05	0	>0.1	No
34	RFLQ	0.98 ± 0.07	0.41 ± 0.08	0.58e	<0.001	Partial
36	QFLQ	1.03 ± 0.05	1.05 ± 0.05	0	>0.1	No
35	RFAR	0.94 ± 0.03	0.91 ± 0.05	0	>0.1	No
40	RFVR	0.95 ± 0.05	1.03 ± 0.07	0	>0.1	No
33	RFRL	1.02 ± 0.04	0.96 ± 0.04	0	>0.1	No

[a] Amino acid sequence of RXLR2 in wild-type and mutants. RFLR is the wild-type. Altered residues are underlined.

[b] Ratio of blue spots in the presence of various RXLR2 mutants of Avr1b-1, compared to the control empty vector when bombarded onto leaves from rps plants (Williams) or Rps1b plants (L77-1863). Averages and standard errors are from 16 pairs of shots.

[c] Ablation calculated as 1- (Rps1b ratio)/(rps ratio) for ratios significantly different between rps and Rps1b.

[d] p values comparing results from rps and Rps1b cultivars were calculated using the Wilcoxon rank sum test.

[e] Ablations for KFLR and RFLQ were significantly different than wildtype (RFLR) with $p < 0.001$.

This example demonstrates that 1) both the RXLR2 and dEER motifs of Avr1b are required for this protein to confer avirulence on *P. sojae* transformants (summarized in Figure 5);

2) the RXLR2 and dEER motifs are not required to trigger an interaction with the *Rps1b* gene product when the Avr1b protein is synthesized in the soybean cytoplasm; 3) when Avr1b protein is directed to be secreted out of the soybean cell, the RXLR2 and dEER motifs are once more required for the protein to trigger an interaction with Rps1b, which is consistent with the motifs being required for the Avr1b protein to re-enter the soybean cell across the plasma cell membrane; 4) fusion of the RXLR and dEER region to GFP (SEQ ID NO: 46) enabled the isolated fusion protein to enter soybean root cells in the absence of the pathogen, but only if the RXLR and dEER motifs were both intact; and 5) RXLR-dependent entry of Avr1b does not require the presence of the pathogen. These observations lead to the conclusion that the RXLR and dEER motifs do indeed have the function of transporting avirulence proteins into host cells.

In addition, the data presented in this Example characterizes the RXLR2 and dEER motifs as follows: 1) arginine at position 1 and leucine at position 3 are essential for function of the RXLR motif. However, there is not a strong requirement for the arginine at position 4. Therefore by functional assays, the oomycete RXLR motif resembles the Plasmodium motif (RxLxD/E/Q) even more closely than previously noted; and 2) the amino acid sequences flanking the RXLR2 and dEER motifs are required in addition to the motifs themselves for the transit of Avr1b into soybean. Further, the region from residues 33 to 71 (19aa to the left of RXLR2 and 6aa to the right of dEER) were sufficient for protein translocation.

The Avr1b protein requires not only the RXLR motif itself, but also non-random surrounding sequences including the dEER motif. These surrounding sequences are not enriched in positive and hydrophobic residues, but instead are enriched in acidic and hydrophilic residues. Furthermore, our RXLR mutagenesis results show that the presence of basic and hydrophobic residues is not sufficient for RXLR function; instead the order of the amino acid residues is very important, and very subtle mutations such as RFLR → RFVR or QFLR abolish function. Therefore, oomycete effectors may utilize a novel mechanism for translocation across the membrane, possibly involving host cell surface machinery (such as a receptor) that is

more complex than just the phospholipid bilayer. The Plasmodium Pexel/VTF motif also requires surrounding sequences that are enriched in acidic and hydrophilic residues and is functionally interchangeable with the oomycete RXLR domain in both erythrocytes and in soybean tissue (this study). Thus oomycetes and Plasmodium both may target host cell surface machinery that is common to plants and vertebrate animals but different than that targeted by animal PTDs. The targeted machinery, if common, must not only be very ancient, but also must serve an irreplaceable function in the host organisms since it must have been preserved against strong negative selection pressure resulting from exploitation by the pathogens.

These results do not indicate which specific flanking sequences are required. However, HMMs constructed from the 10 amino acid residues flanking the upstream and downstream sides of all *P. sojae* and *P. ramorum* Avh RXLR motifs, clearly separated the RXLR motifs of functional avirulence proteins from RXLR motifs obtained by chance from real or permuted proteins sequences. These findings indicate that reliable bioinformatic searches for RXLR effector candidates should include the use of HMMs to evaluate the sequences flanking putative RXLR and dEER motifs.
METHODS: Plasmids and oligonucleotides used in the study are depicted in tabular form as Figure 11A and B and Figure 12A and B, respectively.
P. sojae isolates and transformation: P. sojae isolate P7076 (Race 19) was routinely grown and maintained on V8 agar). The *P. sojae* transformation procedure was described by Dou et al (Dou, D., Kale, S.D., Wang, X., Chen, Y., Wang, Q., Wang, X., Jiang, R.H.Y., Arredondo, F.D., Anderson, R., Thakur, P., McDowell, J., Wang, Y., and Tyler, B.M. (2008) Plant Cell 20(4), 1118-1133).
Characterization of P. sojae transformants: *P. sojae* transformants were selected that grew well on V8 medium with 50 μg/ml G418, and were cultured in V8 liquid medium for 3 days. The mycelia were harvested, frozen in liquid nitrogen and ground to a powder for DNA or RNA extraction. Genomic DNA was isolated from mycelium using known techniques. DNA samples were quantified using a Nanodrop ND-1000 spectrophotometer (Thermo Scientific). The presence of Avr1b-1 transgenes was verified by PCR amplification from 100 ng genomic DNA using a program of 94°C for

2 min, 30 cycles of 94 °C for 30 s, 56 °C for 30 s, 72°C for 30 s, and 72 °C for 5 min with primers of HamF and HamR (TS1). All the transformed *P. sojae* were double-checked by Pst I restriction and/or sequence. RNA was extracted from each sample using RNeasy Plant Mini Kit (QIAGEN, cat # 74904) with β-mercaptoethanol added buffer RLT and genomic DNA was removed using RNase-Free DNase (QIAGEN, cat #79254) according to the manufacturer's recommendations. RNA was quantified using a Nanodrop ND-1000 spectrophotometer. Avr1b-1 transgene transcription was verified by RT-PCR using the internal primers, Avr1bReF and Avr1bReR (TS1) and *P. sojae* actin was used as the reference.

Phenotypic assays for avirulence: Avr1b phenotypic expression was assayed using soybean cultivars HARO(1-7) (rps), Haro13 (Harosoy background, Rps1b), Williams (rps) and L77-1863 (Williams background, Rps1b). Seedlings were grown in the greenhouse or in a growth chamber (Percival AR-36L) with a program of 24°C at daytime and 22°C at night with a 14 hr day length under fluorescent light (250 μmol photons s-1 m-2). The virulence of each transformant was evaluated using hypocotyl inoculation. 1-2 days after the first primary leaf appeared, the hypocotyl of the soybean was wounded with a short incision and the incision was inoculated with a small piece of V8 agar cut from the edge of a 3 day old colony. Thereafter, the plants were incubated in a growth chamber under the conditions described above. The numbers of dead and surviving plants were counted 4 days after inoculation, and summed over 2-5 replicates. The differences between the numbers of surviving plants from rps and Rps1b cultivars were compared using Fisher's exact test. Only the transformants producing a significant difference between rps and Rps1b cultivars were judged as avirulent.

Particle bombardment assays: Particle bombardment assays were carried out using a double-barreled extension of the Bio-Rad He/1000 Particle Delivery System ((Dou, D., Kale, S.D., Wang, X., Chen, Y., Wang, Q., Wang, X., Jiang, R.H.Y., Arredondo, F.D., Anderson, R., Thakur, P., McDowell, J., Wang, Y., and Tyler, B.M. (2008) Plant Cell 20(4), 1118-1133). Analyzing the bombardment data as a ratio between the test and control shots improves the reproducibility of the measurements greatly. The avirulence

activity of the Avr1b-1 constructs was measured as the reduction in the number of blue spots comparing the Avr1b-1 + GUS bombardment with the GUS + control bombardment. For each paired shot the logarithm of the ratio of the spot numbers of Avr1b-1 to that of the control was calculated, then the log-ratios obtained from the Rps1b and non-Rps1b leaves were compared using the Wilcoxon rank sum test.

Bombardment assays of onion bulb cells with GFP constructs: Preparation of DNA-particle mixtures was as described above. 5 mm hemispherical layers of yellow and white onion bulbs were bombarded without the double barrel attachment under a 26 psi vacuum, using a rupture pressure of 1100 psi. The onion layers were incubated between 24-48 hr at 25°C, then viewed with a Zeiss Axioskop2 Plus microscope using a 480nm filter for GFP fluorescence. Images were captured using a Qimaging Retiga 1300 Camera. To further confirm the GFP had been secreted out of the onion cells, plasmolysis was performed for 15 min in 0.8 M mannitol and cells were observed in a Zeiss LSM510 laser scanning confocal microscope (Jena, Germany) with an argon laser excitation wavelength of 488 nm.

RXLR-GFP Fusion Protein Expression and Purification: Residues 33 to 71 of Avr1b (VESPDLVRRSLRNGDIAGGRFLRAHEEDDAGERTFSVTD (SEQ ID NO: 46) including the RXLR1, RXLR2 and dEER motifs were fused to GFP, replacing the Arg9 encoding sequences in vector pR9GFP (SEQ ID NO: 112), called pR9 by Chang et al., (Chang, M., Chou, J.C. and Lee, H.J. (2005) Plant and Cell Physiology 46, 482–488). pR9GFP, which also adds an N-terminal His6 tag, was derived by Chang et al (2005) from Ptat-HA. C43(DE3) *E.coli* cells containing RXLR-GFP fusion constructs or pR9 were grown in 200 mL of LB containing ampicillin 100µg/mL in a 1L baffled flask shaken at 240 rpm at 37°C until reaching an OD of 0.4, at which point the cells were induced by addition of 1mL of 1M IPTG (final [5mM]). After 4 hours further growth at the same conditions, the cells were harvested by centrifugation at 4°C and then stored at -20°C. Visual confirmation of GFP expression was noted by the green color of the bacterial cell pellet. To extract the GFP fusion proteins, cells were thawed on ice for 20 min then 4 mL of lysis buffer (50mM NaH_2P0_4, 300mM NaCl, 10mM imidazole, pH 8.0) were added per 1g of wet cell weight. Lysozyme (Sigma-Aldrich,

cat# L6876) was added to a final concentration of 1mg/mL then the suspension was incubated for 20 min on ice. Sonication (Branson sonifier 150D, with Double stepped micro tip, 3mm) was done at 300W at 15 sec bursts four times with 15 sec cooling periods between each burst. The lysate was centrifuged at 10,000x g for 30 minutes at 4°C, then the supernatant was transferred to a fresh tube and kept on ice until use. 5μL of each sample was stored for SDS-PAGE analysis. Protein purification using Ni-NTA affinity chromatography was performed using the QiaExpressionist protocol. 2mL of 50% Ni-NTA super flow slurry (Qiagen) was loaded on a column. The column was washed twice with 5mL of wash buffer (50mM NaH_2P0_4, 300mM NaCl, 20mM imidazole, pH 8.0). The protein sample was loaded onto the column and then the column was washed twice with 10 vol (10mL) of wash buffer. The protein was eluted with 4 mL of elution buffer (50mM NaH_2P0_4, 300mM NaCl, 200mM imidazole, pH 8.0) into 1mL fractions. These fractions were pooled and concentrated to 300μl using a centrifugal protein concentrator (Amicon Centriplus Centrifugal Filter Device MWCO-3kDa) at 13,500x g. The sample was then mixed with an equal volume of 50mM MES buffer pH 5.8. The protein concentration was measured at 280nm using a nanodrop spectrophotometer (ND-1000) and adjusted to 8mg/mL. All purified GFP preparations fluoresced normally under UV illumination.

RXLR-GFP Fusion Protein Root Cell Transduction Assay: Root tips were cut into lengths of between 0.5 cm and 1 cm, and then were washed with water. Each root tip was completely submerged in 20 μL of the protein solution (8 mg/ml in 25 mM MES pH 5.8) in a eppendorf tube. The samples were incubated overnight at 28°C (~12 hours). The roots were then washed in 200 mL of water for 4 hours while shaken at 100 rpm on a rotary shaker. The roots were then viewed using a Zeiss LSM510 laser scanning confocal microscope with an argon laser excitation wavelength of 488 nm. For nuclear staining, the roots were stained with DAPI (4',6-diamidino-2-phenylindole) (Sigma-Aldrich cat# D8417) and viewed with a 405 nm filter.

Hidden Markof Model Analysis: By using the program HMMER 2.3.2 (Eddy, S.R. (1998). Profile hidden Markov models. Bioinformatics 14, 755-763; and website located at hmmer.janelia.org), an HMM was built from the full set of 765 high quality

candidate effectors identified from the *P. sojae* and *P. ramorum* genomes by Jiang et al (Jiang, R. H. Y., Tripathy, S., Govers, F. & Tyler, B. M. Proc. Natl. Acad. Sci. USA 105, 4874-4879 (2008), using the 10 amino acids on the left side of each RXLR motif together with the 10 amino acids on the right side each RXLR motif. The same procedure was used to build an HMM from a curated list of 191 high quality candidate effectors from *Hyaloperonospora parasitica* developed at the *H. parasitica* genome annotation jamboree in August 2007 and available at pmgn.vbi.vt.edu. To estimate the significance of HMM scores, all proteins (1240) with a predicted N-terminal signal peptide (SP) and the string RXLR located between 30 and 60 amino acids after the SP cleavage site were obtained by translating the genome sequences of *P. sojae* and *P. ramorum* in all reading frames. The sequences of all the putative secreted proteins were permuted (other than the signal peptide) and RXLR-containing sequences were again identified; 639 of the permuted proteins had RXLR strings, indicating that about 639 of the 1240 detected RXLR motifs could be expected by chance. The distributions of HMM scores the set of 1240 real proteins, the 639 permuted proteins and the 765 curated proteins were then calculated. The frequency that a permuted protein received a score between 0 and 5.0 was 0.044. The frequency that a permuted protein received a score better than 5.0 was 0.018.

Accession Numbers: The sequences reported herein have been deposited in the GenBank database, namely Hp Avh341 (EF681127). Accession numbers for sequences already in GenBank are Ps Avr1b-1 (AAM20936), Ps Avr4/6 (ABS50087), Pi Avr3a (CAI72345); Hp Atr1 (AY842877), Hp Atr13 (AY785301).

EXAMPLE 2. Effector host-targeting signals of eukaryotic pathogens bind phosphoinositides or phosphatidic acid

Pathogens of both plants and animals produce effectors and/or toxins that act within the cytoplasm of host cells to suppress host defenses and cause disease. Effector proteins of oomycete plant pathogens utilize N-terminal motifs, RXLR and dEER, to enter host cells, and a similar motif, Pexel (RxLxE/D/Q), is used by Plasmodium effectors to enter erythrocytes. This Example shows that effectors of fungal plant pathogens contain functional variants of the RXLR and dEER motifs, and

that the oomycete and fungal RXLR and dEER motifs, as well as the Plasmodium Pexel motifs, are responsible for binding of the effectors to phosphatidyl-inositol-3-phosphate (PI-3-P) and/or phosphatidyl-inositol-4-phosphate (PI-4-P). Stimulation of host cell entry by PI-4-P, and inhibition by inositol 1,4 diphosphate suggest that phosphoinositide binding mediates cell entry. All the effectors could also enter human cells, suggesting that phosphoinositide-mediated effector entry may be very widespread in plant, animal and human pathogenesis.

Oomycete RXLR and dEER domain binds phosphoinositides

The RXLR and dEER domain of *P. sojae* Avr1b enables translocation of green fluorescent protein (GFP) into plant cells without any pathogen-encoded machinery (see Example 1), and the same is true for two additional bioinformatically predicted effectors, Avh5 (SEQ ID NO: 129) and Avh331 (SEQ ID NO: 127). In these experiments, accumulation of the GFP fusion proteins inside the cells was confirmed by the accumulation of GFP within the nuclei of the cells (a natural property of GFP), and by plasmolysis experiments. In principle, a cell entry domain could bind either a (glyco)protein or (glyco)lipid receptor. After noting that beta-type phosphatidylinositol-4-phosphate kinases from rice and Arabidopsis contained a PI-4-P binding domain consisting of 14 and 11 tandem RXLR and dEER motifs respectively, experiments were conducted to test whether oomycete the RXLR and dEER domain could bind phosphoinositides. An array of 8 different lipids found in plant cell membranes were spotted in decreasing amounts onto a Hybond-C extra membrane. Then the membrane was probed with GFP fused to the N-terminal RXLR and dEER domains of Avr1b (SEQ ID NO: 46), Avh331 (SEQ ID NO: 116) or Avh5 (SEQ ID NO: 113). Figure 6 shows that the Avr1b- and Avh331-GFP fusions bound to PI-4-P while Avh331- and Avh5-GFP fusions bound to PI-3-P. Alanine substitutions mutations in either the RXLR or the dEER motif of any of the three fusions abolished binding, just as they abolished entry into soybean root cells. Fusions of full-length Avh5 (SEQ ID NO: 129) or Avh331 (SEQ ID NO: 127) proteins at their N-termini to glutathione-S-transferase (GST) could also bind the same phosphoinositides as just their N-terminal domains fused at their C-termini to GFP and binding by the full

length proteins also required intact RXLR and dEER motifs (Figure 6) (full-length Avr1b could not be produced in *E. coli*).

To independently confirm binding of the effector RXLR and dEER domains to the phosphoinositides, the binding of the fusion proteins to liposomes composed of phosphatidyl-choline (PC) and phosphatidyl-ethanolamine (PE) was tested. In the absence of phosphoinositides, neither the effector N terminus-GFP fusion proteins nor the full length GST-effector fusion proteins bound the liposomes (Figure 6). However, when either PI-3-P or PI-4-P were included, all the fusion proteins bound to the liposomes. In every case, when any of the RXLR or the dEER motifs were mutated by alanine substitutions, the mutant fusion proteins lost their ability to bind the liposomes (Figure 6).

Identification of fungal effector translocation domains

To test whether fungal effectors contain N-terminal cell entry domains, N-terminal segments from the fungal effectors AvrL567 (SEQ ID NO: 138) and AvrM (SEQ ID NO: 142) of *M. lini* and from AvrPi-ta (SEQ ID NO: 143) of *M. oryzae* were fused to the C-terminus of Avr1b, in the presence of the Avr1b secretory leader, then tested the fusions in a particle bombardment cell re-entry assay that measures the ability of a motif to carry an Avr1b reporter protein back into soybean leaf cells after secretion. Figure 7A shows that all three fungal N-terminal segments had significant ability to deliver Avr1b back into soybean leaf cells.

Since the fungal effectors contained no obvious RXLR or dEER motifs, we decided to define experimentally the range of residues within the RXLR motif of Avr1b that could permit cell entry, using the particle bombardment cell re-entry assay. The results revealed that lysine (K) or histidine (H) but not glutamine (Q) could replace the arginine at position 1 in the motif, that any large hydrophobic residue (isoleucine, I; methionine, M; phenylalanine, F; tyrosine, Y) could replace the leucine at position 3, albeit with varying efficiencies, but valine (V) and alanine (A) could not. At position 4, all residues tested (lysine, K; glutamine, Q; glycine, G) allowed function. Furthermore, the presence of either an L or M residue at position 2 could substitute for a large hydrophobic residue at position 3.

Using this information, one, seven and four potential cell entry motifs were identified in N-terminal regions of AvrL567 (SEQ ID NO: 138), AvrM (SEQ ID NO: 142) and AvrPi-ta (SEQ ID NO: 143), respectively (Figure 7B). The single motif in AvrL567, RFYR, had a particularly good match to the oomycete RXLR motif, and RFYR had already been shown to be functional (Figures 7B and 7C). Four of the candidate motifs in the AvrPi-ta N-terminus, including two that overlapped at one residue (RFLK and KLIFK (SEQ ID NO: 146)), were tested for cell entry activity by substituting them for the RXLR motif of Avr1b. The two single motifs and the overlapping pair were all active in the cell re-entry assay (Figure 7B).

The N-terminus of AvrL567 was subjected to further analysis by mutagenesis and root cell entry assays. Alanine substitutions in the RFYR motif and in two downstream acidic residues that might act as a dEER motif (Figure 7B) (SEQ ID NO: 139), abolished the activity of the AvrL567 N-terminal domain in the particle bombardment cell re-entry assay (Figure 7A and B). To confirm the cell entry activity of the AvrL567 N-terminal domain, it was fused to GFP (creating AvrL567(N) –GFP (SEQ ID NO: 119)), with and without (SEQ ID NO: 120) the alanine substitutions, and then the fusion proteins were tested for cell entry in the soybean root uptake assay. The GFP-fusion with the intact AvrL567 N-terminus (SEQ ID NO: 119) efficiently accumulated in the root cells, including the nuclei, whereas the fusion with the mutated RFYR and acidic residues (rfyr-de-) (SEQ ID NO: 120) did not (not shown). Thus the RFYR motif and the two downstream acidic residues appear to act as a RXLR and dEER motif in *M. lini* AvrL567.

Fungal and apicomplexan effectors bind phosphoinositides

Both filter binding and liposome binding were used to test whether the N-terminal domain of AvrL567 bound phosphoinositides. AvrL567(N)-GFP (SEQ ID NO: 119) bound PI-3-P in both assays. Binding of AvrL567(N)-GFP to PI-4-P was also be detected in the liposome assay though it is not as strong as to PI-3-P. Mutation of the RXLR and dEER-like motif to alanines (rfyr-de- mutant) (SEQ ID NO: 120) resulted in a loss of binding to the phosphoinositides in both assays.

Filter binding assays were used to determine if five additional fungal effectors could bind phospholipids. The N-terminal sequences of the following effectors were fused to GFP: *Magnaporthe grisea* AvrPita (SEQ ID NO: 224), *Puccinia graminis* Ps87 (SEQ ID NO: 226); *Melampsora lini* AvrM (SEQ ID NO: 228); *Melampsora lini* AvrP123 (SEQ ID NO: 230); *Melampsora lini* AvrP4 (SEQ ID NO: 232). The results, summarized in tabular form in Figure 15, showed that all five effectors bound phosphatidic acid. Mutations in RXLR-like motifs found in AvrPtia (SEQ ID NO: 225), Ps87 (SEQ ID NO: 227), AvrM (SEQ ID NO: 228), AvrP123 (SEQ ID NO: 230) and AvrP4 (SEQ ID NO: 232) all abolished binding to phosphatidic acid. In the case of AvrPita, the mutant protein (SEQ ID NO: 225) was unable to enter soybean root cells, whereas the wildtype protein (SEQ ID NO: 224) could enter root cells, suggesting that binding of AvrPita to phosphatidic acid was required to enter plant cells.

The host targeting signals (HTS) of three *Plasmodium falciparum* effectors, PfGBP (SEQ ID NO: 121), PfHRPII (SEQ ID NO: 123), and Pf1615c (SEQ ID NO: 125) can carry Avr1b into soybean leaf cells and onion bulb epidermal cells8. The three signals can also carry purified GFP into soybean root cells and this activity requires intact Pexel motifs. To test whether the three signals also could bind phosphoinositides, the HTS-GFP fusion proteins were tested using filter binding and liposome binding assays. The PfGBP HTS fusion (SEQ ID NO: 121) could bind PI-4-P and also, more weakly, PI-3-P (Figure 8A). The PfHRPII HTS fusion (SEQ ID NO: 123) could bind PI-3-P, and also rather weakly, PI-4-P (Figure 8B) The Pf1615c HTS fusion (SEQ ID NO: 125) could bind specifically to PI-3-P (Figure 8C). Liposome binding assays confirmed binding of all the fusion proteins to PI-3-P or PI-4-P (Figures 8D-F). In both assays, alanine substitutions in the Pexel motifs of each effector abolished phosphoinositide binding (Figure 8A-F) (SEQ ID NO: 122, 124, 126).

Modulation of effector entry by exogenous phosphoinositides

The binding of phosphoinositides to the effector cell entry domains suggested that these phospholipids might serve as a cell entry receptor in each case. Tomato cells

secreted PI-4-P when stimulated by fungal xylanase, suggesting that free PI-4-P might exist in the plant apoplast24. Therefore, increasing the concentration of free phosphoinositide by exogenous addition might stimulate RXLR and dEER-mediated uptake. To test this hypothesis, a soluble form of PI-4-P, di-octanoyl-PI-4-P (250 μM), was mixed with the Avr1b GFP fusion, Avr1b(N)-GFP, for 30 min prior to exposure to soybean roots. After 9 hr, strong stimulation of Avr1b(N)-GFP uptake by PI-4-P was evident (Figure 9A). The phospholipids, PI, PC and PE all did not stimulate uptake. Uptake of the fungal AvrL567(N)-GFP fusion (SEQ ID NO: 119) was also strongly stimulated by PI-4-P (Figure 9B), even though it binds most strongly to PI-3-P.

A synthetic cell entry motif composed of nine-arginine residues (Arg9) (SEQ ID NO: 112) was previously shown to deliver Avr1b into soybean leaf cells and into onion epidermal leaf cells in particle bombardment cell re-entry assays. The motif could also enable uptake of purified GFP into soybean root cells8 and into maize and onion cells. The mechanism of uptake has been proposed to be a plant form of macropinocytosis. The Arg9-GFP fusion protein binds PI-3-P, PI-4-P and phosphatidyl serine, albeit weakly (Figure 9D). Figure 9C shows that di-octanoyl-PI-4-P does not stimulate uptake of the Arg9-GFP fusion protein in soybean root cells, suggesting that the stimulation by PI-4-P is specific to RXLR and dEER-mediated uptake. This conclusion is supported by the observation that exogenous PI-4-P did not promote the uptake of Avr1b(N)-GFP (SEQ ID NO: 46) and AvrL567(N)-GFP (SEQ ID NO: 119) proteins containing alanine substitutions in the RXLR and dEER motifs.

Inositol-1,4-diphosphate (IP2) represents the hydrophilic head-group of PI-4-P. Preincubation with 100 μM IP2 inhibited binding of Avr1b(N)-GFP (SEQ ID NO: 46) to PI-4-P-containing liposomes and could completely block binding of AvrL567(N)-GFP (SEQ ID NO: 119) to PI-4-P-containing liposomes, presumably via competitive inhibition. To test whether IP2 could block effector uptake in planta, which would imply that a PI-4-P-like molecule mediated uptake in planta, Avr1b(N)-GFP (SEQ ID NO: 46) or AvrL567(N)-GFP (SEQ ID NO: 119) was preincubated with 500 μM IP2 for 30 min prior to exposure to soybean roots. IP2 almost completely blocked uptake of both Avr1b(N)-GFP (Figure 9A) (SEQ ID NO: 46) or AvrL567(N)-GFP (Figure 9B)

(SEQ ID NO: 119) into soybean cells. IP2 could not inhibit the binding of Arg9-GFP to liposomes (Figure 9D) and uptake of Arg9-GFP (SEQ ID NO: 112) was completely unaffected by preincubation with IP2 (Figure 9C), supporting the conclusion that IP2 specifically blocks RXLR and dEER motif-mediated protein uptake.

Effector entry into human cells

Phosphatidyl-inositol-phosphates are universally found in eukaryotic cells. Since a number of human and animal diseases are caused by fungi and oomycetes, as well as by apicomplexan parasites, we tested the possibility that RXLR and dEER motifs might mediate protein entry into human cells, using the human lung epithelial cell line A549 as a model. Avr1b(N)-GFP (Figure 10A) (SEQ ID NO: 46), AvrL567(N)-GFP (Figure 10B) (SEQ ID NO: 119) and PfHRPII(N)-GFP (Figure 10C) (SEQ ID NO: 123) (SEQ ID NO: 60) could all enter the A549 cells, but entry did not occur if alanine substitutions were present in the RXLR or Pexel motifs of the proteins (SEQ ID NO: 47,48,120,124). In these experiments, accumulation of the GFP fusion proteins inside the cells was confirmed by the accumulation of GFP within vesicle-like structures within the cells, and by the fact that the cells were treated with protease (trypsin) prior to photographing. Protein accumulation was strongly inhibited in each case by 1,4-IP2, supporting the hypothesis that entry was mediated by phosphoinositide binding. Exogenous di-octanoyl-PI-4-P did not stimulate accumulation, suggesting that the availability of PI-4-P or other phosphatidylinositides in the growth medium was not limiting. Inositol 1,3 diphosphate (1,3-IP2), the headgroup of PI-3-P, could also inhibit entry of PfHRPII(N)-GFP (Figure 10C) (SEQ ID NO: 123), consistent with the observation that the protein bound PI-3-P more strongly than PI-4-P in the filter-binding assay.

Discussion: Two independent assays, namely filter-binding and liposome-binding, demonstrated that the N-terminus of all seven effectors tested could bind to either PI-3-P or PI-4-P, but not to PI-5-P nor to any other phospholipids tested. The primary structures of the RXLR and dEER effector domains do not resemble any known phosphoinositide binding domains. However, the binding of the pathogen RXLR and dEER domains to phosphoinositides is concordant with the binding of the RXLR and

dEER domains of rice and Arabidopsis β-type PI-4-kinases. In the oomycete effector proteins, the dEER motif is variably spaced from the RXLR motif, so if residues from both motifs contact the phosphoinositide head group, the protein must fold so as to bring the two motifs into proximity. The three dimensional structure of the RXLR and dEER or Pexel domain is not yet available for any oomycete or apicomplexan effector proteins, respectively. However, the crystal structure of AvrL567 has been determined. In this structure, the RFYR motif adopts a beta-stranded conformation on the surface of the protein. It will be interesting to determine if the structure of AvrL567 changes in solution in the presence of a phosphoinositide.

The stimulation of protein entry into soybean root cells by PI-4-P and the inhibition of entry by IP2 together support the hypothesis that binding to phosphoinositides mediates entry of these pathogen effectors into plant cells. Similar findings with a human lung epithelial cell line suggest the possibility that effectors of oomycetes and fungi that infect humans and other animals might enter host cells via a similar mechanism. This mechanism appears to be different than reported for other peptides with cell entry activity because entry of those other peptides is not dependent on phosphoinositides.

These data demonstrate that several fungal effectors contain N-terminal domains that are capable of carrying Avr1b into soybean leaf cells. Within these domains are RXLR-like motifs that can replace the Avr1b RXLR motif in carrying reporter proteins into host cells. In the one case tested in detail so far, AvrL567, the motif RFYR was identified as necessary for the activity of the N-terminal cell entry domain, and for the binding of the domain to PI-3-P. The mutagenesis survey of the Avr1b RXLR motif and the diverse functional motifs found in the fungal effectors together suggest that a wide diversity of RXLR-like sequences support binding of phosphoinositides. Bioinformatic screens with the highly redundant motif suggested by the data identifies huge numbers of matches, most of which are likely spurious as judged by searches of permuted protein sequences. Thus, it is very likely that there are additional requirements for phosphoinositide binding. Both the RXLR and dEER

motifs of Avr1b and the Pexel motifs of Plasmodium effectors are insufficient by themselves to facilitate cell entry; in both cases flanking sequences are required.

Other than in apicomplexan parasites, eukaryotic pathogens of humans and other animals have not been reported to produce effector proteins that can cross host membranes into the cytoplasm of host cells. The finding that phosphoinositide-binding effectors from oomycete and fungal plant pathogens can cross the membranes of human cells predicts that oomycete and fungal pathogens of humans and other animals may also utilize this mechanism to debilitate their hosts. Possible examples include oomycete pathogens of marine animals from the genera *Saprolegnia* and *Aphanomyces*, extracellular fungal pathogens such as *Pneumocystis carinii*, *Coccidioides immitus* and *Aspergillus fumigatus*, and intra-phagosomal fungal pathogens of humans such as *Cryptococcus neoformans* and *Histoplasma capsulatum*.

The binding of phosphoinositides or other polar lipids to effector cell entry domains from diverse kingdoms will provide a powerful biochemical tool for screening or directly isolating new candidate effector proteins from all classes of microbes. It may also enable detection of phosphoinositide-binding plant proteins (or other polar-lipid-binding proteins) that can traffic through the apoplast and enter into target cells to transduce signals. Some precedents for such proteins already exist, such as the *Drosophila antennapedia* transcription factor that can move from cell to cell via an arginine-rich cell entry motif.

Understanding the role of phosphoinositides and other polar lipids in pathogen effector entry also opens the possibility of targeting cell entry domains for preventative or therapeutic intervention in both agriculture and medicine. The finding that IP2 can block effector entry into both plant and human cells provides a proof-of-concept for this approach.

METHODS SUMMARY

Cloning was performed according to known molecular biology techniques. Proteins were expressed in *E.coli* BL21DE3 and purified using affinity chromatography. Plasmids and oligonucleotides used in the study are depicted in tabular form as Figures 13A-F and Figures 14A-D, respectively.

Soybean seeds were germinated in vermiculite for 3-5 days. Roots were washed with water thoroughly to remove any debris. Approximately 1.5 cm root tips were cut and placed into the protein solution (50 µl- 25 mM MES pH 5.8, 50 µg protein) and incubated for 12-15 hr at 28°C. Then the root tips were rinsed with water and washed in 75 mL of water for 2 hr on an orbital shaker at 90 rpm. Roots were examined using a Zeiss LSM510 laser scanning confocal microscope with an argon laser excitation wavelength of 488 nm.

Lipid filter arrays were prepared by pipetting 1 µL PI-3-P, PI-5-P (Cayman Chemical), PS, PC, PE, PA, or PI-4-P (Avanti Polar Lipids, Cayman Chemical) at various concentrations on Hybond-C extra membranes.

Liposomes were prepared from a suspension of 0.71 µg/ml phosphatidyl-choline, 0.29 µg/ml phosphatidyl-ethanolamine (PC/PE) or 0.64 µg/ml phosphatidyl-choline, 0.26 µg/ml phosphatidyl-ethanolamine, 0.1 µg/ml phosphatidyl-inositol-phosphate (PC/PE/PI-x-P). The lipid mixtures were dried under vacuum overnight, then the resultant lipid films were rehydrated at 1 mg/mL (total lipid) in 20 mM Tris-HCl (pH 6.8) 100 mM NaCl, 2 mM dithiothreitol by three cycles of freeze-thawing. Large unilamellar vesicles were formed by extruding the lipid suspension through a 0.1- µM filter (nucleopore track-etch membrane, Whatman) 20 times and were used immediately. Effector fusion proteins were centrifuged at 100,000 g for 20 min at 25°C prior to assay to remove protein aggregates. 10 µg of protein was added to 50 µg of liposomes and incubated for 1 hr at room temperature. Protein-liposome mixtures were centrifuged at 100,000 g for 15 min at 25°C. Pellets containing liposome-bound proteins and supernatants containing free proteins, were then analyzed by SDS-PAGE.

EXAMPLE 3. Assay for screening compound libraries to identify novel compounds that interfere with the RXLR and dEER-mediated uptake of effector proteins into plant or human cells

The binding of phosphoinositides (PI-3-P or PI-4-P) and phosphatidic acid to effector cell entry domains indicates that these phospholipids may serve as a cell entry receptors. Increasing the concentration of free phosphoinositide such as di-octanoyl-PI-4-P by exogenous addition stimulated RXLR and dEER-mediated uptake of the

Avr1b GFP fusion, Avr1b(N)-GFP, into soybean roots and human cells. Furthermore, preincubation with inositol 1,4 diphosphate (IP2) inhibited binding of Avr1b(N)-GFP to PI-4-P-containing liposomes presumably via competitive inhibition. In addition, IP2 almost completely blocked uptake of both Avr1b(N)-GFP into soybean root cells and human cells in cell culture. Therefore, an assay is devised for screening compound libraries to identify novel compounds that interfere with the RXLR and dEER-mediated uptake of effector proteins into plant or human cells, through inhibition of the binding of, or interaction between phospholipids PI-3-P or PI-4-P and RXLR and dEER motif containing proteins. Plasmid encoding the Avr1b sequence are expressed in BL21 *E. coli* cells and the protein are purified and diluted into appropriate binding buffer at an appropriate concentration, and thirty microliters are dispensed into each well coated 96 or 384 well plates using an automated dispenser. Through a robotized transfer mechanism involving steel pins, each of the Avr1b protein-containing wells (in a 96 or 384 well plate) receive 300 nanoliter of a compound from the compound libraries, followed by incubating the plate at room temperature for 60 minutes. An equal volume of 2x stock solution of fluorescently labeled soluble PI-4-P (Echelon Inc. BODIPY FL Phosphatidylinositol(4) Phosphate catalog #C-04F6a; BF-PI-4-P) is prepared in suitable buffer and 30 microliter of this solution is dispensed into each well of Avr1b coated preincubated 384 well plates using an automated dispenser. After BF-PI-4-P addition, the plate is incubated in dark for 60 minutes, followed by the measurement of fluorescence, utilizing a Synergy plate reader integrated with a biostack. The reactions are performed in duplicates and with negative controls, where the interactions are measured in the absence of the protein or fluorescently labeled BF-PI-4-P, and positive controls where the interaction is measured in the presence of a range of concentrations of inositol 1,4 diphosphate (IP2). The readouts are stored and analyzed for the identification of potential inhibitors of the reaction. Statistical analysis are performed utilizing a combination of parameters and compounds that showed statistically significant inhibition are selected. Briefly, the background absorbance is subtracted from the test reads. Subsequently, the net absorbance is compared to controls wells, that did not receive the test compounds and the percent

decrease in absorbance is measured by the following formula: Percent inhibition = [(Fluorescence in test well/Fluorescence in control wells) x 100]. In excess of one hundred thousand drug-like, diverse heterocyclic chemical compounds are screened during this process for their potential to inhibit the interactions between BF-PI-4-P and Avr1b. These compounds are obtained from several sources including established chemical vendors like Asinex, Analyticon, Biomol, Bionet, ChemDiv, Enamine, Maybridge, Spectrum, TimTec as well as a range of diversity oriented synthesis compounds that have been generated by academic research laboratories from around the world. Typical screening identifies several hundred compounds that inhibit the reaction at a statistically significant >40% levels. Successful events in this initial screen lead to the consolidation of select wells from the original library stock to generate a new second generation of plate for screening the activity of these compounds at three compound concentrations to allow the calculation of a preliminary IC50 value. A select group of compounds is then selected that showed >50% inhibition. Larger quantities of select compounds are ordered from the specific vendors (above) for rescreening in the soybean root or human cell uptake assays for their potential to inhibit the uptake of Avr1b(N)-GFP into soybean roots cells or human cells in culture.

EXAMPLE 4. Screening assays for novel compounds that inhibit plant oomycete or fungal infection through blocking of RXLR and dEER containing effector protein action

To further characterize candidate protective compounds obtained from the RXLR and dEER protein screen, the ability of the compounds to protect against oomycete or fungal pathogen infection are tested in detached leaf assays. The following detached leaf assays are used for soybean, potato, tomato, tobacco, grape, rice, and wheat. The assays are used to test for infection by *Phytophthora* oomycete pathogens (soybean, potato, tomato, tobacco), downy mildew oomycete pathogens (tobacco and grape), rust fungi (soybean and wheat), *Magnaporthe* blast fungi (rice and wheat), and powdery mildew fungi (soybean, potato, tomato, tobacco, grape, wheat).

Expanded leaves are removed from young growth chamber-grown plants with the petioles intact (soybean, potato, tomato, tobacco, grape), or are clipped from the mother plant with sterile scissors (wheat and rice). The petioles or cut ends of the leaves are placed into plastic test tubes containing an aqueous solution of a suitable range of concentrations of each compound (determined from the biochemical IC_{50}). The leaves are then fastened into a horizontal position, but with the petioles or cut ends bent down into the tubes. The plants are then placed in a lighted growth chamber at 30% humidity for 6 hr to enable the compounds to be drawn into the leaves by transpiration.

The plants are then inoculated with pathogen spores. Phytophthora infections (*P. sojae* on soybean; *P. infestans* on tomato and potato; *P. parasitica* on tobacco) are initiated by spraying the leaves with an aqueous suspension of zoospores at a suitable concentration. Infections with rust fungi (*Phakopsora pachyrizi* for soybean; *Puccinia striiformis* f. sp. tritici (stripe rust) *Puccinia triticina* (leaf rust), *Puccinia graminis* f. sp. tritici (stem rust) for wheat) are initiated by spraying the leaves with an aqueous suspension of urediniospores at a suitable concentration. Infections with downy mildew oomycetes (*Peronsopora tabacina* on tobacco; *Plasmopora viticola* on grape) and infections of Magnaporthe blast fungi (*Magnaporthe oryzae* for rice; *Magnaporthe grisea* on wheat) are initiated by spraying the leaves with an aqueous suspension of conidia at a suitable concentration. In each case, after spraying with the pathogen spore suspension, the plants are replaced into the growth chamber at high (90%) humidity at a suitable temperature (15°C for *Phytophthora infestans* and downy mildew oomycetes; 25°C for P. sojae; 20°C for all others) until symptoms develop (3-7 days). Infections with powdery mildew (*Microsphaera diffusa* on soybean; *Erysiphe cichoracearum* on potato and tobacco; *Leveillula taurica* on tomato or potato; *Erysiphe necator* on grape; *Blumeria graminis* f. sp. *tritici* on wheat) are done dry. The arrays of plant leaves are placed into a dusting tower and heavily infected leaves are introduced into the top of the tower and shaken vigorously for one minute, then the spores are allowed to settle for 20 min.

The plants are replaced into growth chambers maintained at 70% humidity (except *Microsphaera diffusa* which is favored by low humidity of 30%) and 25°C.

To assay disease development in each case, the leaves are photographed, and from the photographs the numbers of lesions on each leaf are counted (total lesions and sporulating lesion) and the areas of the lesions are determined digitally. The results are assessed statistically by reference to negative controls (water), positive controls (benomyl for the fungi and metalaxyl for the oomycetes).

EXAMPLE 5. Protecting plants against oomycete infection by blocking effector entry using peptide receptor mimics

Eukaryotic pathogens such as oomycetes, fungi and apicomplexan parasites deliver hundreds of effector proteins into the cytoplasm of their host cells. Delivery of these proteins is key to the pathogenic success of these organisms. The similarity between oomycete and apicomplexan effector delivery systems has been noted for some time. The discovery that inositol 1,4 diphosphate can inhibit oomycete and fungal effector uptake (Example 2) shows that effector entry can be blocked by externally applied small molecular weight compounds. This Example describes experiments that test whether infection by oomycetes, and possibly by fungi, can be mitigated by inhibiting effector entry using host-synthesized peptides that mimic inositol 1,4 diphosphate. Biotrophic and hemi-biotrophic oomycete pathogens that are likely to use RXLR and dEER effectors include more than 80 species of *Phytophthora* and more than 500 species of downy mildews that together attack almost every crop species and horticultural species of economic importance. Peptides that could inhibit RXLR and dEER effector entry could thus provide broad-spectrum protection against many of these pathogens. Even if protection is narrow, and multiple peptides must be selected for each species of pathogen, this approach offers an important new weapon against these highly adaptable pathogens. The fact that one fungal effector from a rust pathogen also may use phosphoinositides to enter host cells suggests that the protection provided by anti-RXLR and dEER peptides may extend to many biotrophic and hemi-biotrophic fungal pathogens such as rusts, smuts, powdery mildews and the rice blast fungus.

At least three commercially available phage display libraries are screened against a panel of effectors that have been well-characterized and/or are strongly expressed at the outset of *P. sojae* or *P. infestans* infection. The phage are eluted from the effectors using a rising concentration gradient of IP2 or soluble PI-P in order to identify those phage that have the greatest affinity for the PI-P binding sites of the effectors. The candidates obtained are evaluated for their binding to all panel members, and to RXLR and dEER mutants of the panel members. In addition, their affinity for both soluble and liposome-bound PI-4-P and PI-3-P is measured. Synthetic peptides corresponding to candidates with the highest affinities are prepared commercially and tested for their ability to inhibit uptake into plant and human cells. The most promising peptides at this point (broadest specificity, highest affinity and/or strongest inhibition of effector entry) are targeted for optimization of their breadth and affinity of binding. Two optimization strategies are used. Firstly, PCR-directed random mutagenesis of selected peptides is carried out and high affinity, broad-spectrum mutants are selected by phage display. Loss-of-activity mutants are also characterized to help identify important residues. As a parallel alternative approach, targeted mutagenesis of selected peptides is carried out based on bioinformatic analysis of all the phage peptide sequences obtained (both high quality and low quality peptides). Surface Plasmon resonance and NMR characterization of the binding of the most promising peptides to their target effector(s) also yields important information. The cycle of selecting promising peptides and optimizing them is repeated as needed, or until little further improvement is obtained. At this point the most promising peptides are fused to selected secreted plant proteins, and the chimeric proteins are expressed in plant tissues by transient expression. The expression levels and stabilities of the chimeric proteins are assessed, as well as the ability of the expressed proteins to reduce entry of effector proteins and reduce infection by *P. sojae* or *P. infestans*. The resistance of the plant tissue to additional pathogens is also evaluated. Stable transgenic plants expressing the chimeric proteins are produced, and are evaluated systematically for disease resistance against diverse pathogens.

Initially the Ph.D.-C7C random peptide phage display library available from New England Biolabs, Inc. is screened. In this library, a loop of 7 random amino acid residues is constrained by a disulfide bond at the base of the loop where it is fused to the N-terminus of the pIII coat protein. This configuration was chosen because the loop will eventually be transferred to a secreted plant protein, and the disulfide bond will ensure that the loop has a similar structure in that context as on the phage. The library contains 1.2×10^9 independent phage, providing about 60% statistical coverage of the total theoretical complexity of a heptapeptide library ($207 = 1.3 \times 10^9$). Other possible libraries for screening include Ph.D.-7 and the Ph.D.-12 libraries that contain 7 or 12 random residues respectively, but without a disulfide bond; both have a complexity of around 2.8×10^9.

Two approaches are used to select RXLR and dEER-specific phage. In the first, individual effectors are screened. Since this allows interactions with residues outside the immediate RXLR and dEER region, inhibitory peptides with a narrow specificity are likely to be obtained. In a complementary approach, the phage are selected on several different effectors successively, in order to only obtain peptides with broad specificity. Both strategies are adjusted as needed.

The one-effector-at-a-time strategy targets Avr1b, Avh331 (Avr1k), Avh5, Avh6 and Avh172. Because the first two effectors are avirulence gene products that trigger plant defense responses mediated by resistance (R) genes, the efficacy of candidate inhibitory peptides is also tested in planta by their ability to inhibit the R gene mediated response to the effectors. Avh6 and Avh172 are major early-expressed effectors, so targeting them singly also has a measurable effect on pathogen virulence. Avh5 is included because its NMR characterization is well advanced. In each case, two different fusions are produced: GFP and GST (glutathione-S-transferase) to reduce the chance of selecting phage that bind to an irrelevant part of the protein.

In the second strategy, three pools of effectors are created, and the phage are successively selected on the different pools. One example of such a set of pools is: pool 1 = Avr1b + Avh331 + Avh6; pool 2 = Avh5 + Avh172 + Avh152; pool 3 = Avh38 + Avh260 + AvrL567. Each effector listed is either an avirulence protein or a

strongly-early-expressed *P. sojae* effector, or is otherwise well-characterized. By using pools, the risk that a single chosen effector may be problematic is reduced, and by using three different pools for the successive selection steps, the likelihood of finding broad specificity peptides is increased. The order of the pools used for selection is varied. The composition of the pools may be varied once data on the specificity of each effector for PI-3-P or PI-4-P is available, and/or if production of some chosen effector proteins in *E. coli* proves problematic.

Panning is carried out in microtiter tray wells; if sufficient enrichment of peptides is not seen in the wells, then the proteins are bound onto beads and the beads are used for panning. The phage are step eluted with different concentrations of inositol diphosphates (IP2) or soluble (e.g. di-hexanoyl) phosphatidyl phosphates. The choice of 1,3 IP2, 1,4 IP2, PI-3-P or PI-4-P, and the choice of concentrations is finalized once more precise data on the binding constant of the effectors for the phosphoinositides is available.

Characterization of discovered peptides for binding to multiple effectors, and for ability to block entry of key oomycete and fungal effectors into plant cells.

Each selected peptide is screened against a panel of all the effectors mentioned listed above, plus a selection of 10 *P. infestans* infection-induced effectors and several fungal effectors. RXLR and dEER region mutants are included to identify peptides that interact with those motifs. Phage with the broad specificity and a set of peptides with complementary sets of targets are identified. Screening is done in a western dot blot format in which effectors bound to a filter are probed with the phage and then with an anti-M13 antibody. Alternatively, the phage are panned against effectors arrayed in microtiter wells, and then detected by spotting onto an *E. coli* lawn with a replicator. The affinity of the phage for the effector is initially estimated by doing binding experiments in the presence of different concentrations of PI-Ps or IP2s. An oomycete effector protein microarray containing all 1440 effectors from *P. sojae*, *P. infestans*, *P. ramorum* and *H. arabidopsidis* is ideal for comprehensive screening of the most promising phage. The most promising peptides are tested for the ability to block effector-GFP entry into root cells. To obtain sufficient peptides for these

experiments, synthetic peptides are ordered from a commercial supplier. A quantitative cell entry assay using luciferase and suspension culture cells may also be employed. Binding of the most promising peptides to key effectors is further characterized by NMR and surface plasmon resonance (SPR).

Concatenation and mutagenesis of the most promising peptides to further optimize broad-spectrum binding and ability to block effector entry into plant cells.

Bioinformatic comparisons of peptide sequences having different affinities and ranges of specificity provide important starting clues about the potential for using targeted mutations to improve affinity and specificity of the identified peptides. Alternatively, phage display technology is a proven platform for improving binding via random mutagenesis. A single randomized oligonucleotide is used to mutagenize the 21 nucleotides encoding each peptide loop. Selection of phage on a range of different effectors is used to improve the breadth of specificity. Selection of phage in the presence of free peptide having the original sequence is used to select for improved affinity. An alternative approach to improving the breadth of specificity is to concatenate several peptides, with a spacer or linker sequence in between. This is an acceptable construction for in planta expression. The concatenated peptides are tested to ensure that they retain their original affinities and breadth of specificity.

Fuse peptides to small secreted plant proteins, and test the effects of their expression in planta on effector entry and on disease resistance.

For evaluation *in planta,* the peptide mimics are fused to larger proteins normally produced during infection to promote the peptides' stability and reduce their potential susceptibility to endogenous plant proteases. The fusions are evaluated in three steps: (i) exogenous application of purified proteins to plant tissues; (ii) transient expression in plants; and (iii) expression in stable transgenic plants. A variety of candidate proteins are evaluated for fusions with the peptide mimics, including highly stable plant proteins such as PR1a, lipid transfer proteins, protease inhibitors and proteases. Fusion to a protease inhibitor promotes stability, while conversely, fusion to a protease more effectively targets pathogen effectors for proteolysis. Generally, the mimic is attached to the C-terminus of the "carrier" protein

via a suitable spacer so that the native N-terminal secretory leader can be used. Initially a single peptide mimic is attached to each carrier. Once attachment of single peptides has been validated, multiple peptides are attached in tandem to improve the breadth of binding and/or for better efficacy against effectors with multiple phophoinositide binding sites. C-terminal green fluorescent protein (GFP) fusions are used to evaluate the stability and localization of the proteins *in planta*.

Expression in E. coli or Pichia pastoris and evaluation of purified proteins.

Fusion proteins are expressed in *E. coli* or, due to the necessity to correctly form disulfide bonds, in eukaryotic expression system based on *Pichia pastoris*. The purified peptide-fusion proteins are tested for effector binding *in vitro* to ensure they retain binding activity as fusions. They are then introduced into leaf and root tissues (by infiltration and direct uptake, respectively) from soybean and *N. benthamiana* to test their stability *in planta* (via western blots) and to test their ability to inhibit the uptake of exogenously applied effector-reporter fusions into the plant cells. Uptake assays based on suspension cultures cells and on protoplasts may also be used to distinguish between stability and effectiveness in effector uptake inhibition.

Transient expression in planta.

Excellent virus-based transient expression systems now exist for soybean and Arabidopsis. Infiltration of *Agrobacterium tumefaciens* strains harboring vectors designed to deliver gene expression constructs into plant cells locally (Agroinfiltration) and particle bombardment have also been used extensively. Initially a quantitative "double-barrel" particle bombardment assay is used to measure the ability of plant-expressed peptide fusion proteins to interfere with effector entry into soybean cells, using either the native effectors, Avr1b or Avh331 (i.e. Avr1k), or fusions of other effectors to an Avr1b reporter. By using GFP-fusions in conjunction with onion epidermal cell bombardment direct visualization of the localization of the peptide fusions or the targeted effector (or both together if one carries a red fluorescent protein, e.g mCherry) is possible. Bi-molecular fluorescence complementation (BiFC; "split-YFP") is used to verify effector-peptide interaction in planta in the onion system.

In order to evaluate the potential effect of peptide-fusion expression in planta on pathogen infection, the BPMV system is used to transiently express the peptide fusion proteins in soybean and Agroinfiltration to transiently express the proteins in *N. benthamiana*. Versions that include GFP to facilitate evaluation of stability and localization are used. Transcription of the constructs is confirmed using RT-PCR or northern analysis. Protein levels are evaluated by western blots, and confocal microscopy is used to verify that the proteins are being delivered to the apoplast.

Areas of plant tissue transiently expressing the peptide fusions are inoculated with *P. sojae* (soybean) or *P. infestans* (*N. benthamiana*) and disease development is evaluated. Empty vectors and vectors with the entire construct minus the peptide mimic are used as negative controls in these experiments. *N. benthamiana* tissue is tested for its response to the blue mold downy mildew pathogen *Peronospora tabacina*. Soybean leaf tissue is tested for resistance to the soybean rust fungus, *Phakopsora pachyrhizi*.

Hairy root cultures of soybean expressing the peptides are created, and assayed them for *P. sojae* resistance, to show that expression of Avr1b and Avh331 confers increased susceptibility to *P. sojae*. Stably transformed soybean, *N. benthamiana* and *Arabidopsis* plants expressing the fusion proteins are tested and display resistance to a variety of oomycete and fungal pathogens.

CLAIMS

We claim:

1. A method of inhibiting entry of a pathogenic effector protein into a cell, comprising the step of

binding a blocking compound to one or more motifs of a pathogenic effector protein, wherein said one or more motifs are bound by a polar lipid as a prerequisite to translocation of the effector protein into a host cell.

2. The method of claim 1 wherein said one or more motifs are selected from the group consisting of RxLR, dEER and Pexel motifs.

3. The method of claim 1 wherein said one or more motifs is an RxLR motif, and said RxLR motif is followed by at least one aspartate or at least one glutamate residue within a 60 amino acid carboxy terminal amino acid sequence.

4. The method of claim 1 wherein said one or more motifs is an RxLR motif, and said RxLR motif includes at least one of a two or three amino acid sequence selected from the group consisting of: arginine, any amino acid, leucine; histidine, any amino acid, leucine; lysine, any amino acid, leucine; arginine, any amino acid, isoleucine; histidine, any amino acid, isoleucine; lysine, any amino acid, isoleucine; arginine, any amino acid, methionine; histidine, any amino acid, methionine; lysine, any amino acid, methionine; arginine, any amino acid, tyrosine; histidine, any amino acid, tyrosine; lysine, any amino acid, tyrosine; arginine, any amino acid, phenylalanine; histidine, any amino acid, phenylalanine; lysine, any amino acid, phenylalanine; arginine, any amino acid, tryptophan; histidine, any amino acid, tryptophan; lysine, any amino acid, tryptophan; arginine, any amino acid, valine; histidine, any amino acid, valine; lysine, any amino acid, valine; arginine, leucine; histidine, leucine; lysine, leucine; arginine, isoleucine; histidine, isoleucine; lysine, isoleucine; arginine, methionine; histidine, methionine; lysine, methionine; arginine, tyrosine; histidine, tyrosine; lysine, tyrosine;

arginine, phenylalanine; histidine, phenylalanine; lysine, phenylalanine; arginine, tryptophan; histidine, tryptophan; lysine, tryptophan; arginine, valine; histidine, valine; and lysine, valine.

5. The method of claim 1 wherein said polar lipid is selected from phosphoinositides, phospholipids, and sphingolipids.

6. The method of claim 5 wherein said phosphoinositide is selected from the group consisting of phosphatidyl-inositol-3-phosphate (PI-3-P), phosphatidyl-inositol-4-phosphate (PI-4-P), phosphatidyl-inositol-5-phosphate (PI-5-P), phosphatidyl-inositol-3,4-diphosphate (PI-3,4-P2), phosphatidyl-inositol-3,5-diphosphate (PI-3,5-P2), phosphatidyl-inositol-4,5-diphosphate (PI-4,5-P2), phosphatidyl-inositol-3,4,5-triphosphate (PI-3,4,5-P3), lysophosphatidyl-inositol-3-phosphate (LPI-3-P), lysophosphatidyl-inositol-4-phosphate (LPI-4-P), lysophosphatidyl-inositol-5-phosphate (LPI-5-P), lysophosphatidyl-inositol-3,4-diphosphate (LPI-3,4-P2), lysophosphatidyl-inositol-3,5-diphosphate (LPI-3,5-P2), lysophosphatidyl-inositol-4,5-diphosphate (LPI-4,5-P2), and lysophosphatidyl-inositol-3,4,5-triphosphate (LPI-3,4,5-P3), and phosphatidyl-inositol (PI), and lysophosphatidyl-inositol (LPI).

7. The method of claim 1 wherein said polar lipid is selected from the group consisting of phosphatidyl-serine (PS), phosphatidyl-glycerol (PG), phosphatidyl-ethanolamine (PE), phosphatidyl-choline (PC), lysophosphatidyl-serine (LPS), lysophosphatidyl-glycerol (LPG), lysophosphatidyl-ethanolamine (LPE), lysophosphatidyl-choline (LPC), phosphatidic acid (PA), lysophosphatidic acid (LPA), sphingosine-1-phosphate (S-1-P), ceramide-1-phosphate (C-1-P), a glycosylphosphatidylinositol (GPI)-protein anchor, a glycosylsphingosylinositol (GSI)-protein anchor, a glycosyl phosphoryl inositol ceramide (GPIC) and sphingomyelin (SM).

8. The method of claim 1, wherein said blocking compound is a synthetic peptide.

9. The method of claim 1, wherein said blocking compound is selected from the group consisting of: an inositol phosphate, an inositol sulfate, an inositol carboxylate, an inositol arsenate, an inositol phosphorothioate, a hexose phosphate, a hexose sulfate, a hexose carboxylate, a hexose arsenate, a hexose phosphorothioate, a hexitol phosphate, a hexitol sulfate, a hexitol carboxylate, a hexitol arsenate, a hexitol phosphorothioate, a polyol phosphate, a polyol sulfate, a polyol carboxylate, a polyol arsenate, a polyol phosphorothioate, a phosphorylated glycan, a sulfated glycan, a carboxylated glycan, a glycan arsenate, or a glycan phosphorothioate.

10. The method of claim 1, wherein said cell is a plant cell.

11. The method of claim 9, wherein said plant cell is of a type selected from the group consisting of wheat, maize, rice, sorghum, barley, oats, millet, soybean, common bean (*Phaseolus* species), green pea (*Pisum* species), cowpea, chickpea, alfalfa, clover, tomato, potato, tobacco, pepper, egg plant, grape, strawberry, raspberry, cranberry, blueberry, blackberry, hops, walnut, apple, peach, plum, pistachio, apricot, almond, pear, avocado, cacao, coffee, tea, pineapple, passionfruit, coconut, date and oil palm, citrus, orange, lemon, grapefruit, safflower, carrot, sesame, common bean, banana, citrus, papaya, macadamia, guava, pomegranate, pecan, Brassica species, canola, cabbage, cauliflower, mustard, cucurbits, pumpkin, cantalope, squash, zucchini, melon, cotton, sugar cane, sugar beets, sunflower, lettuce, onion, garlic, ornamental cut flowers; and grass.

12. The method of claim 1 wherein said cell is an animal cell.

13. The method of claim 11, wherein said animal cell is from an animal selected from the group consisting of cattle, sheep, pigs, goats, horses, cats, dogs, chickens, turkeys, bees, salmon, trout, bass, catfish, shellfish, crayfish, lobsters, shrimp, and crabs.

14. The method of claim 11 wherein said animal cell is a human cell.

15. The method of claim 11 wherein said animal cell is a red blood cell, a lymphocyte, a macrophage, a neutrophil, a dendritic cell, a spleen cell, a thymus cell, a liver cell, a nerve cell, a brain cell, a lung cell, a muscle cell, or an epithelial cell.

16. The method of claim 1 wherein said pathogenic effector protein is from an oomycete.

17. The method of claim 15 wherein said oomycete is selected from the group consisting of: a *Phytophthora* species, *Phytophthora infestans*, *Phytophthora sojae*, *Phytophthora ramorum*, *Phytophthora parasitica*, *Phytophthora capsici*, *Phytophthora nicotianae*, *Phytophthora cinnamomi*, *Phytophthora cryptogea*, *Phytophthora drechsleri*, *Phytophthora cactorum*, *Phytophthora cambivora*, *Phytophthora citrophthora*, *Phytophthora citricola*, *Phytophthora megasperma*, *Phytophthora palmivora*, *Phytophthora megakarya*, *Phytophthora boehmeriae*, *Phytophthora kernoviae*, *Phytophthora erythroseptica*, *Phytophthora fragariae*, *Phytophthora heveae*, *Phytophthora lateralis*, *Phytophthora syringae*, a *Pythium* species, *Pythium ultimum*, *Pythium aphanidermatum*, *Pythium irregulare*, *Pythium graminicola*, *Pythium arrhenomanes*, *Pythium insidiosum*, a downy mildew species, a *Peronospora* species, *Peronospora tabacina*, *Peronospora destructor*, *Peronospora sparsa*, *Peronospora viciae*, a *Bremia* species, *Bremia lactucae*, a *Plasmopora* species, *Plasmopora viticola*, *Plasmopara halstedii*, a *Pseudoperonospora* species, *Pseudoperonospora cubensis*, *Pseudoperonospora humuli*, a *Sclerospora* species, *Sclerospora graminicola*, a *Peronosclerospora* species, *Peronosclerospora philippinesis*, *Peronosclerospora sorghi*, *Peronosclerospora sacchari*, a *Sclerophthora* species, *Sclerophthora rayssiae*, *Sclerophthora macrospora*, a *Albugo* species, *Albugo candida*, a *Aphanomyces* species, *Aphanomyces cochlioides*, *Aphanomyces euteiches*, *Aphanomyces invadans*, a *Saprolegnia* species, *Saprolegnia parasitica*, and a *Achlya* species.

18. The method of claim 1 wherein said pathogenic effector protein is from a fungus.

19. The method of claim 17 wherein said fungus is selected from the group consisting of: a rust fungus, a smut fungus, a bunt fungus, a powdery mildew fungus, a *Puccinia* species, *Puccinia striiformis*, *Puccinia graminis*, *Puccinia triticina* (syn. *Puccinia recondita*), *Puccinia sorghi*, *Puccinia schedonnardii*, *Puccinia cacabata*, a *Phakopsora* species, *Phakopsora pachyrhizi*, *Phakopsora gossypii*, a *Phoma* species, *Phoma glycinicola*, a *Ascochyta* species, *Ascochyta gossypii*, a *Cryphonectria* species, *Cryphonectria parasitica*, a *Magnaporthe* species, *Magnaporthe oryzae*, a *Gaeumannomyces* species, *Gaeumannomyces graminis*, a *Synchytrium* species, *Synchytrium endobioticum*, a *Ustilago* species, *Ustilago maydis*, *Ustilago tritici*, *Ustilaginoidea virens*, a *Tilletia* species, *Tilletia indica*, *Tilletia caries*, *Tilletia foetida*, *Tilletia barclayana*, a *Erysiphe* species, *Erysiphe necator,* a *Blumeria* species, *Blumeria graminis*, *Podosphaera oxyacanthae*, a *Alternaria* species, *Alternaria alternata*, a *Botrytis* species, *Botrytis cinerea*, a *Diaporthe* species, *Diaporthe phaseolorum*, a *Fusarium* species, *Fusarium graminearum*, *Fusarium oxysporum*, *Fusarium moniliforme*, *Fusarium solani*, a *Leptosphaeria* species, *Leptosphaeria maculans*, *Leptosphaeria maydis*, a *Macrophomina* species, *Macrophomina phaseolina*, a *Monilinia* species, *Monilinia fructicola*, a *Mycosphaerella* species, *Mycosphaerella graminicola*, *Mycosphaerella fijiensis*, *Mycosphaerella tassiana*, *Mycosphaerella zeae-maydis*, a *Phialophora* species, *Phialophora gregata*, a *Phymatotrichopsis* species, *Phymatotrichopsis omnivora*, a *Taphrina* species, *Taphrina deformans*, a *Aspergillus* species, *Aspergillus flavus*, *Aspergillus parasiticus*, *Aspergillus fumigatus*, a *Verticillium* species, *Verticillium dahliae*, *Verticillium albo-atrum*, *Rhizoctonia solani*, *Ophiostoma ulmi*, *Ophiostoma novo-ulmi*, a *Septoria* species, *Septoria avenae*, a *Pyrenophora* species, *Pyrenophora tritici-repentis*, a *Colletotrichum* species, *Colletotrichum graminicola*, a *Sclerotinia* species, *Sclerotinia sclerotiorum*, a *Sclerotium* species, *Sclerotium rolfsii*, a *Thielaviopsis* species, *Thielaviopsis basicola*, a *Coccidioides* species, *Coccidioides immitus*, a

Paracoccidioides species, *Paracoccidioides braziliensis*, a *Pneumocystis* species, *Pneumocystis carinii*, a *Histoplasma* species, *Histoplasma capsulatum*, a *Cryptococcus* species, *Cryptococcus neoformans*, a *Candida* species, *Candida albicans*, a microsporidial species, a *Enterocytozoon* species, a *Encephalitozoon* species and *Encephalitozoon cuniculi*.

20. The method of claim 1, wherein said pathogen effector protein is from a protozoon.

21. The method of claim 19 wherein said protozoon is selected from the group consisting of: an apicomplexan parasite, a *Plasmodium* species, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae*, a *Babesia* species, *Babesia bovis*, *Babesia bigemina*, a *Cryptosporidium* species, *Cryptosporidium parvum*, a *Toxoplasma* species, *Toxoplasma gondii*, a *Trypanosomatid* species, a *Trypanosoma* species, *Trypanosoma brucei*, *Trypanosoma cruzi*, *Trypanosoma congolense*, *Trypanosoma vivax*, a *Leishmania* species, *Leismania donovani*, an amebozoan parasite, an *Entamoeba* species, *Entamoeba histolytica*, a *Mastigamoeba* species, a *Schistosoma* species, a *Onchocerca* species, and a *Giardia* species.

22. The method of claim 1 wherein said step of binding includes the step of providing to a plant or animal a sufficient quantity of said blocking compound so that it is present to bind to said one or more motifs of said pathogenic effector protein prior to entry of said pathogenic effector protein into cells of said plan or animal.

23. A method of identifying pathogenic effector proteins, comprising the step of:

determining whether a protein includes an RxLR amino acid sequence;

comparing, using hidden markov modeling (HMM), a ten amino acid sequence flanking a carboxy terminus of said RxLR amino acid sequence and a ten amino acid sequence flanking an amino terminus of said RxLR amino acid sequence, and,

i) if a score resulting from said HMM comparison is 0, concluding that said RXLR amino acid sequence is not an authentic RxLR motif; or

ii) if a score resulting from said HMM comparison is greater than 5, concluding that said RxLR amino acid sequence is an authentic RxLR motif; or

iii) if a score resulting from said HMM comparison is 5 or less and greater than 0, concluding that said comparison is inconclusive.

24. A method for screening compounds to identify whether they are potential blocking compounds for inhibiting entry of pathogenic effector proteins into a cell, comprising the steps of:

providing one or more proteins each of which has one or more motifs which are bound by polar lipids as a prerequisite to translocation;

exposing a candidate compound to said one or more proteins; and

determining whether said candidate compound binds to said one or more motifs of said one or more proteins, and if binding occurs determining that said compound is a potential blocking compound for inhibiting entry of pathogenic effector proteins into a cell.

25. The method of claim 23 wherein said one or more proteins provided in said providing step are pathogenic effector proteins derived from a bacterial, protozoal, fungal, oomycete or nematode source.

26. A method for predicting whether or not a gene in a genome of a pathogen encodes for a pathogen effector molecules, comprising the step of identifying DNA sequences in a genome which encode proteins with amino acid sequences that satisfy the structural requirements of one or both of an RXLR motif and a dEER motif.

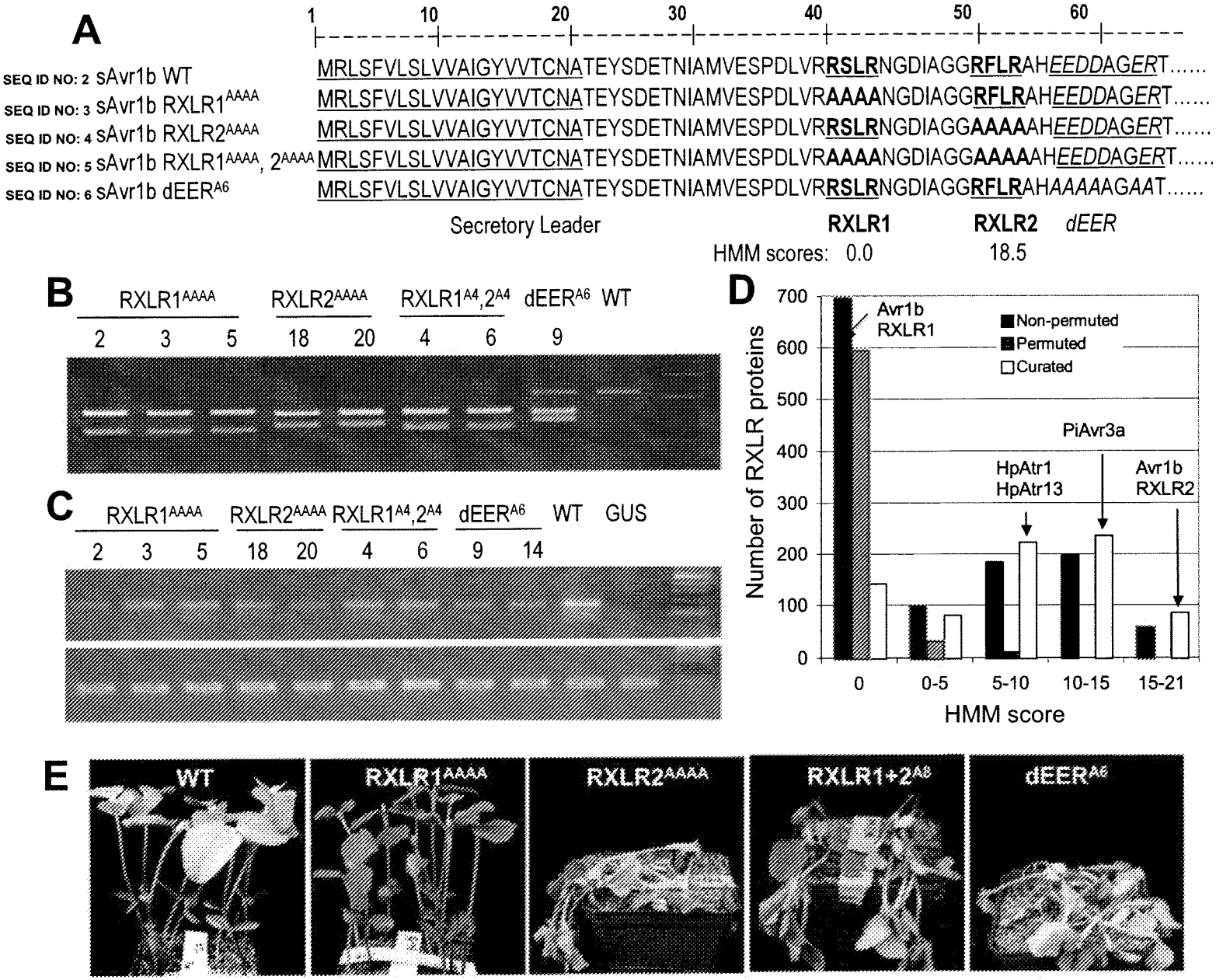

A
1 10 20 30 40 50 60
SEQ ID NO: 2 sAvr1b WT MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRRSLRNGDIAGGRFLRAHEEDDAGERT.....
SEQ ID NO: 3 sAvr1b RXLR1AAAA MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRAAAANGDIAGGRFLRAHEEDDAGERT.....
SEQ ID NO: 4 sAvr1b RXLR2AAAA MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRRSLRNGDIAGGAAAAAHEEDDAGERT.....
SEQ ID NO: 5 sAvr1b RXLR1AAAA, 2AAAA MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRAAAANGDIAGGAAAAAHEEDDAGERT.....
SEQ ID NO: 6 sAvr1b dEERA6 MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRRSLRNGDIAGGRFLRAHAAAAAGAAT.....
Secretory Leader
RXLR1 RXLR2 dEER
HMM scores: 0.0 18.5
B
RXLR1AAAA RXLR2AAAA RXLR1A4,2A4 dEERA6 WT
2 3 5 18 20 4 6 9
C
RXLR1AAAA RXLR2AAAA RXLR1A4,2A4 dEERA6 WT GUS
2 3 5 18 20 4 6 9 14
D
Number of RXLR proteins
700 600 500 400 300 200 100 0
Avr1b RXLR1
Non-permuted
Permuted
Curated
PiAvr3a
HpAtr1 HpAtr13
Avr1b RXLR2
0 0-5 5-10 10-15 15-21
HMM score
E
WT
RXLR1AAAA
RXLR2AAAA
RXLR1+2A8
dEERA6

Figure 1A-E

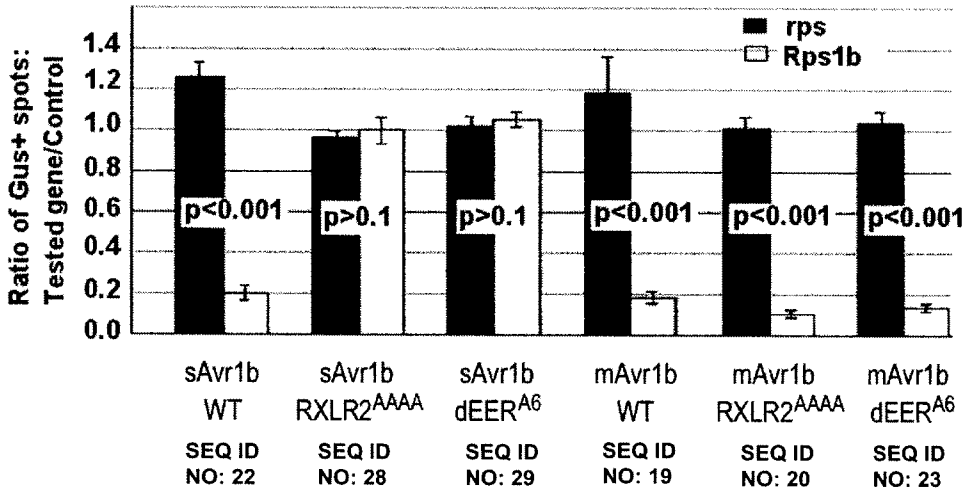

Ratio of Gus+ spots:
Tested gene/Control
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
rps
Rps1b
p<0.001
p>0.1
p>0.1
p<0.001
p<0.001
p<0.001
sAvr1b WT
SEQ ID NO: 22
sAvr1b RXLR2AAAA
SEQ ID NO: 28
sAvr1b dEERA6
SEQ ID NO: 29
mAvr1b WT
SEQ ID NO: 19
mAvr1b RXLR2AAAA
SEQ ID NO: 20
mAvr1b dEERA6
SEQ ID NO: 23

Figure 2

A

SEQ ID NO: 2 sAvr1b WT MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRRSLRNGDIAGGRFLRAHEEDDAGERtfsvt.....

SEQ ID NO: 4 sAvr1bRXLR2^{AAAA} MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRRSLRNGDIAGGAAAAAHEEDDAGERtfsvt.....

SEQ ID NO: 8 PsAvh4/6::Avr1bCt MGLHKGFFVAVALLALLIVAAPADAITDESQPRDATIVDAPLTGRGANARYLRTSTSIIKAPDAQLP-STKAAIASSVTKEEEERKISTGLSKLRQKLSKRFHDIPDWLLQLQAFLSVGLHHLtfsvt.....

SEQ ID NO: 7 HpAvh341::Avr1bCt MRLHYVGPVVAAALAASAHGLQVVSDSASDLQRTDEARQQPYINDKTKRFLTSEDKDLPLLVTS-DGYASLLPQGGDNDRVLRSADIGDGDYEEERSKIKKRHKRKSHGHGVLDtfsvt.....

SEQ ID NO: 9 mAvr1bCt M( )tfsvt.....

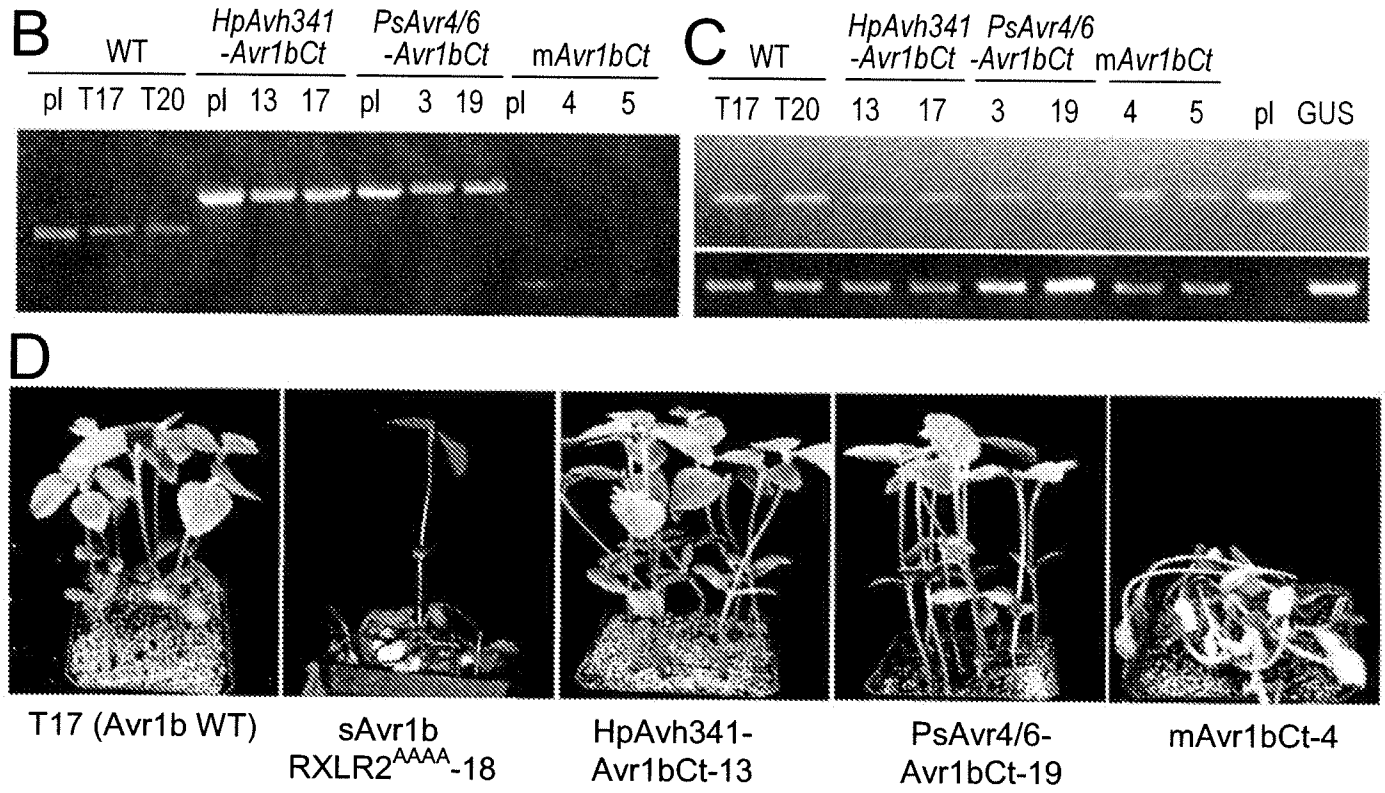


Figure 3 A-D

A

SEQ ID NO: 22 sAvr1b MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRRSLRNGDIAGG-**RFLR**----------AH*EEDD*AG*E*RTFSV••••

SEQ ID NO: 41 sAvr1b RXLR2^{Arg9} MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRRSLRNGDIAGG-RRRRRRRRR---AH*EEDD*AG*E*RTFSV••••

SEQ ID NO: 42 sAvr1b RXLR2^{TAT} MRLSFVLSLVVAIGYVVTCNATEYSDETNIAMVESPDLVRRSLRNGDIAGGYGRKKRRQRRRAH*EEDD*AG*E*RTFSV••••

SEQ ID NO: 43 sAvr1b^{PfGBP} MRLSFVLSLVVAIGYVVTCNA----------DKYEKAVDYGFRES**RILAE**G*ED*TCARK*E*KTTLRKSKQKTFSV••••

SEQ ID NO: 44 sAvr1b^{PfHRP} MRLSFVLSLVVAIGYVVTCNA----------FNNNLCSKNAKGLNLNK**RLLHE**TQAHV*DD*AAHHAHHVA*D*TFSV••••

SEQ ID NO: 45 sAvr1b^{Pf1615c} MRLSFVLSLVVAIGYVVTCNA----------SYNKINSSSTYTHS**RILKQ**LEFITL*EE*KTVNALQ*E*MLDDSVTFSV••••

B

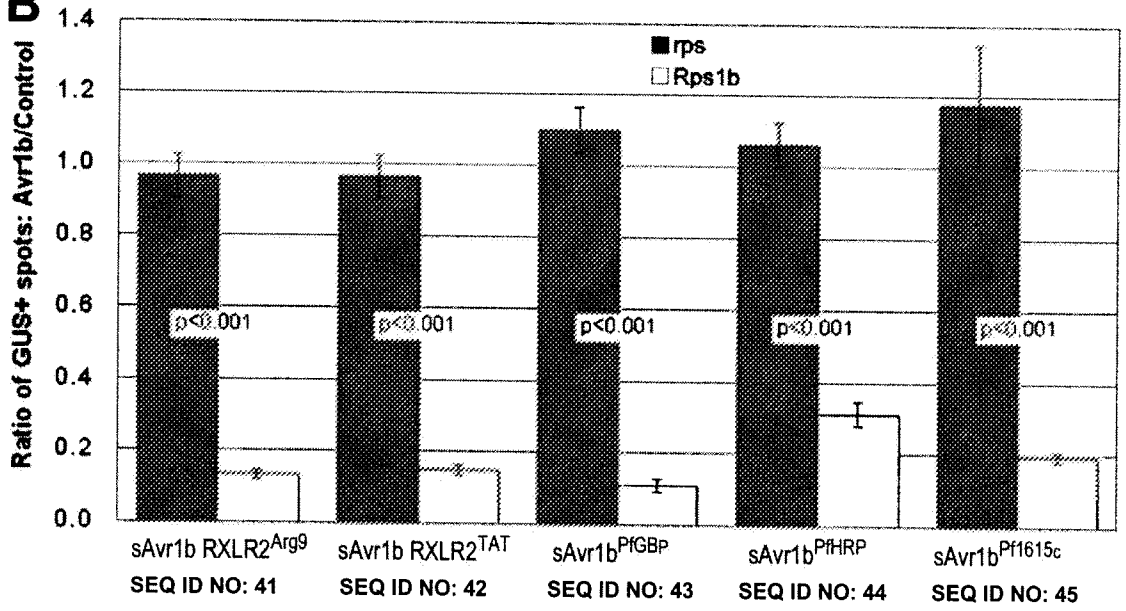


Figure 4 A and B

		Phenotype of *P.sojae* Transformants Rps1b	rps	Cell Death in Bombardment Assay Rps1b	rps
SEQ ID NO: 22	sAvr1b WT	A	V	Y	N
SEQ ID NO: 28	sAvr1b RXLR2^{AAAA}	V	V	N	N
SEQ ID NO: 29	sAvr1b dEER^{A6}	V	V	N	N
SEQ ID NO: 19	mAvr1b WT	NT		Y	N
SEQ ID NO: 20	mAvr1b RXLR2^{AAAA}	NT		Y	N
SEQ ID NO: 21	mAvr1b dEER^{A6}	NT		Y	N
SEQ ID NO: 10	sAvr1b RXLR1^{AAAA}	A	V	NT	
SEQ ID NO: 5	sAvr1b RXLR1^{AAAA},2^{AAAA}	V	V	NT	
SEQ ID NO: 8	- PsAvh171::Avr1bCt	A	V	NT	
SEQ ID NO: 7	-HpAvh341::Avr1bCt				
SEQ ID NO: 33-37, 40	sAvr1b RXLR2	NT		N	N
SEQ ID NO: 38, 39	sAvr1b RXLR2	NT		P	N
SEQ ID NO: 41, 42	sAvr1b RXLR2$^{Arg9/TAT}$	NT		Y	N
SEQ ID NO: 43-45	sAvr1b$^{PfGBP/PfHRP/Pf1615c}$	NT		Y	N

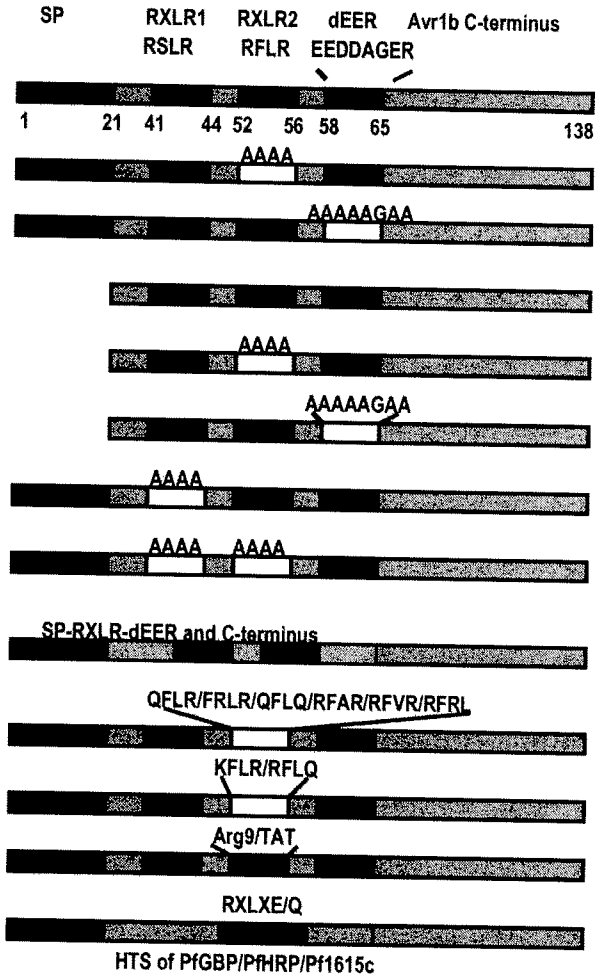


Figure 5

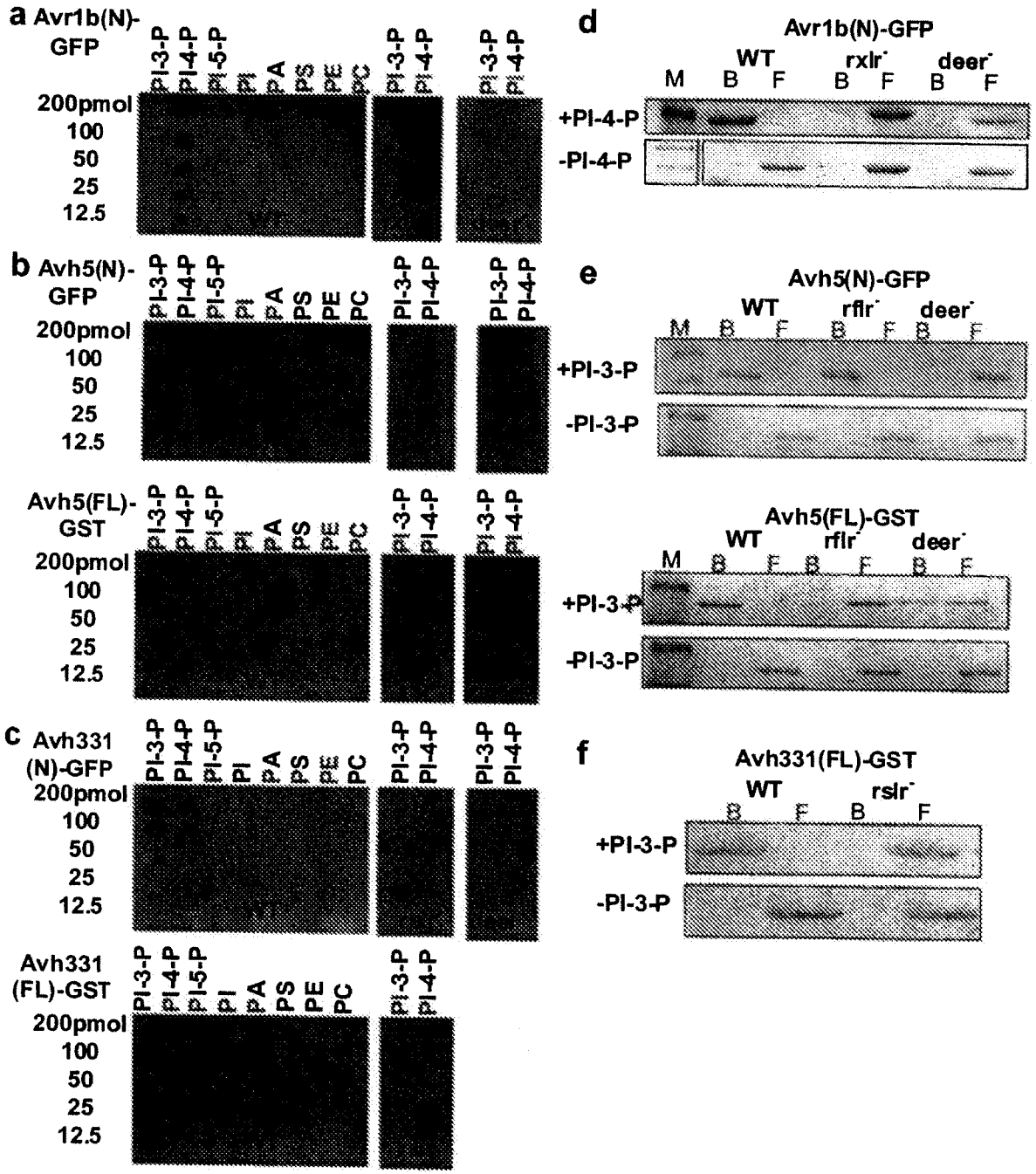

a Avr1b(N)-GFP
PI-3-P
PI-4-P
PI-5-P
PI
PA
PS
PE
PC
200pmol
100
50
25
12.5
b Avh5(N)-GFP
Avh5(FL)-GST
c Avh331 (N)-GFP
Avh331 (FL)-GST
d
Avr1b(N)-GFP
WT
rxlr-
deer-
M
B
F
+PI-4-P
-PI-4-P
e
Avh5(N)-GFP
rflr-
+PI-3-P
-PI-3-P
Avh5(FL)-GST
f
Avh331(FL)-GST
rslr-

Figure 6A-F

G

Protein	Sequence	Binding PI-3-P	Binding PI-4-P	Relative Ablation	Root Cell Entry	Human Cell Entry	Inhibition by IP2
SEQ ID NO: 46 Avr1b	|TEYSDETNIAMVESPDLVRRSLRNGDIAGGRFLRAHEEDDAGERTFSVTD|	-	+	1.00	Y	Y	Y
SEQ ID NO: 47 Avr1b rxlr-	|TEYSDETNIAMVESPDLVRAAAANGDIAGGAAAAAHEEDDAGERTFSVTD|	-	-	0.00	N	N	nt
SEQ ID NO: 48 Avr1b deer-	|TEYSDETNIAMVESPDLVRRSLRNGDIAGGRFLRAHAAAAAGERTFSVTD|	-	-	0.00	N	N	nt
SEQ ID NO: 116 Avh331	|LTCATSEQQTRPELCFFFSVRSSWPSTISDGACLALVSAEQGATAGRNT LSLRSMMATEDMATSTRSLRSQATNVDDDANVSIENRGM|	+	+	1.07	Y	Y	Y
SEQ ID NO: 117 Avh331 rslr-	|LTCATSEQQTRPELCFFFSVRSSWPSTISDGACLALVSAEQGATAGRNT LSLRSMMATEDMATSTAAAASQATNVDDDANVSIENRGM|	-	-	0.00	N	N	nt
SEQ ID NO: 118 Avh331 deer-	|LTCATSEQQTRPELCFFFSVRSSWPSTISDGACLALVSAEQGATAGRNT LSLRSMMATEDMATSTRSLRSQATNVAAAANVSIANAGM|	-	-	0.00	N	N	N
SEQ ID NO: 113 Avh5	|TRVPDDANLQSVNAPVQTVTRSRRFLRTADTDIVYEPKVHNPGK|	+	-	nt	Y	Y	Y
SEQ ID NO: 114 Avh5 rflr-	|TRVPDDANLQSVNAPVQTVTRSRAAAATADTDIVYEPKVHNPGK|	-	-	nt	N	N	nt
SEQ ID NO: 115 Avh5 deer-	|TRVPDDANLQSVNAPVQTVTRSRRFLRTAATAIVYAPAVHNPGK|	-	-	nt	N	N	nt
SEQ ID NO: 119 AvrL567	|MEHVPAELTRVSEGYTRFYRSPTASVILSGLVKVKWDNEQMTMP|	+	+	0.70	Y	Y	Y
SEQ ID NO: 120 AvrL567rfyr-de-	|MEHVPAELTRVSEGYTAAAASPTASVILSGLVKVKWANAQMTMP|	-	-	0.00	N	N	nt
SEQ ID NO: 123 PfHRPII	|FNNNLCSKNAKGLNLNKRLLHETQAHVDDAHHAHHVAD|	+	±	1.15	Y	Y	Y
SEQ ID NO: 124 PfHRPII pexel-	|FNNNLCSKNAKGLNLNKAAAAATQAHVDDAHHAHHVAD|	-	-	nt	N	N	nt
SEQ ID NO: 121 PfGBP	|DKYEKAVDYGFRESRILAEGEDTCARKEKTTLRKSKQK|	±	+	0.97	Y	nt	nt
SEQ ID NO: 122 PfGBP pexel-	|DKYEKAVDYGFRESAAAAAGEDTCARKEKTTLRKSKQK|	-	-	nt	N	nt	nt
SEQ ID NO: 125 Pf1615c	|SYNKINSSSTYTHSRILKQLEFITLEEKTVNALQEMLDDSV|	+	-	1.11	Y	nt	nt
SEQ ID NO: 126 Pf1615c pexel-	|SYNKINSSSTYTHSAAAAALEFITLEEKTVNALQEMLDDSV|	-	-	nt	N	nt	nt
SEQ ID NO: 112 Arg9	|RRRRRRRRR|	±	±	1.13	Y	Y	N

H

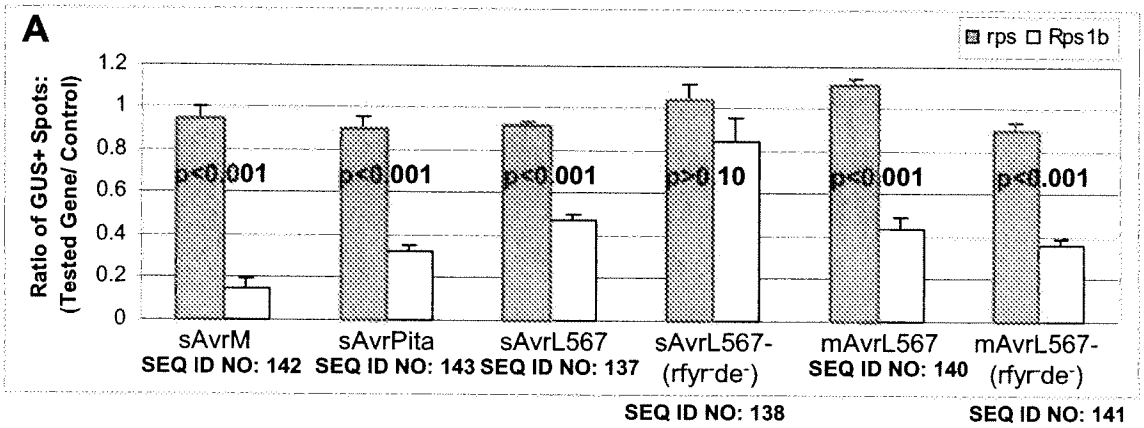
secretion signal peptide

Avr1b- MRLSFVLSLVVAIGYVVTCNA|*TEYSDETNIAMVESPDLVR**RSLR**NGDIAGG**RFLR**AH**EEDD**AGER*|TFSVTDLWNKVAAKKLAKA

6xHis-tag **EK Site** **HA-tag**

Arg$_9$-GFP- MRGS**HHHHHH**GMASMTGGQQMGRNLY**DDDDK**DRWGSRS|***RRRRRRRRR***|EFRSTMSG**YPYDVPDYA**GSMGSGIQRPTST SSLVAAAAT***MSKEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTL***

Figure 6 G-H



B

		Relative Ablation
SEQ ID NO: 22	Avr1b ····TEYSDETNIAMVESPDLVRRSLRNDIAGG.....RFLR.....AHEEDDAGERTFSV····	1.00
SEQ ID NO: 149	Avr1b(AvrL567) ····----------.....RFYR.....----------····	0.84
SEQ ID NO: 144	Avr1b(Pita1) ····----------.....KRRAYI.....----------····	0.51
SEQ ID NO: 145	Avr1b(Pita2) ····----------.....RAALK.....----------····	0.41
SEQ ID NO: 146	Avr1b(Pita3,4) ····----------RFLKLIFKTDSTDIQ----------····	0.56
SEQ ID NO: 143	AvrPita ····HPVYDYNPIPNHIHGDLKRRAYIERYSQCSDSQASEIRAALK SCAELASWGYHAVKNDNRLFKLIFKTDSTDIQN····	0.76
SEQ ID NO: 142	AvrM ····NNLGTVPDVPHQIPNDKSGTPAIEDPKDNKGFNKALKSTPESEKLGTSSVEGIP QPEFDRGFLRPFGAKMRFLKPDQVQKLSTQDLITYMAEKDKNVR····	0.85
SEQ ID NO: 137	AvrL567 ····MEHVPAELTRVSEGYTRFYRSPTASVILSGLVKVKWDNEQMTM····	0.70
SEQ ID NO: 138	AvrL567(rfyr-de-) ····MEHVPAELTRVSEGYTAAASPTASVILSGLVKVKWANAQMTM····	0.00

C

			Relative Ablation
SEQ ID NO: 1	Avr1b	TEYSDETNIAMVESPDLVRRSLRNDIAGGRFLRAHEEDDAGERTFSV····	1.00
SEQ ID NO: 3	Avr1b(AAAA)	----------AAAA----------····	0.00
SEQ ID NO: 38	Avr1b(1K)	----------K----------····	0.42
SEQ ID NO: 147	Avr1b(1H)	----------H----------····	0.89
SEQ ID NO: 37	Avr1b(1Q)	----------Q----------····	0.00
SEQ ID NO: 148	Avr1b(3F)	----------F----------····	1.01
SEQ ID NO: 149	Avr1b(3Y)	----------Y----------····	0.84
SEQ ID NO: 150	Avr1b(3I)	----------I----------····	0.52
SEQ ID NO: 151	Avr1b(3M)	----------M----------····	0.89
SEQ ID NO: 40	Avr1b(3V)	----------V----------····	0.00
SEQ ID NO: 35	Avr1b(3A)	----------A----------····	0.00
SEQ ID NO: 39	Avr1b(4Q)	----------Q----------····	0.74
SEQ ID NO: 152	Avr1b(4G)	----------G----------····	0.88
SEQ ID NO: 153	Avr1b(2M3V)	----------MV----------····	0.71
SEQ ID NO: 154	Avr1b(2L3T4G)	----------LTG----------····	0.56

Figure 7A-C

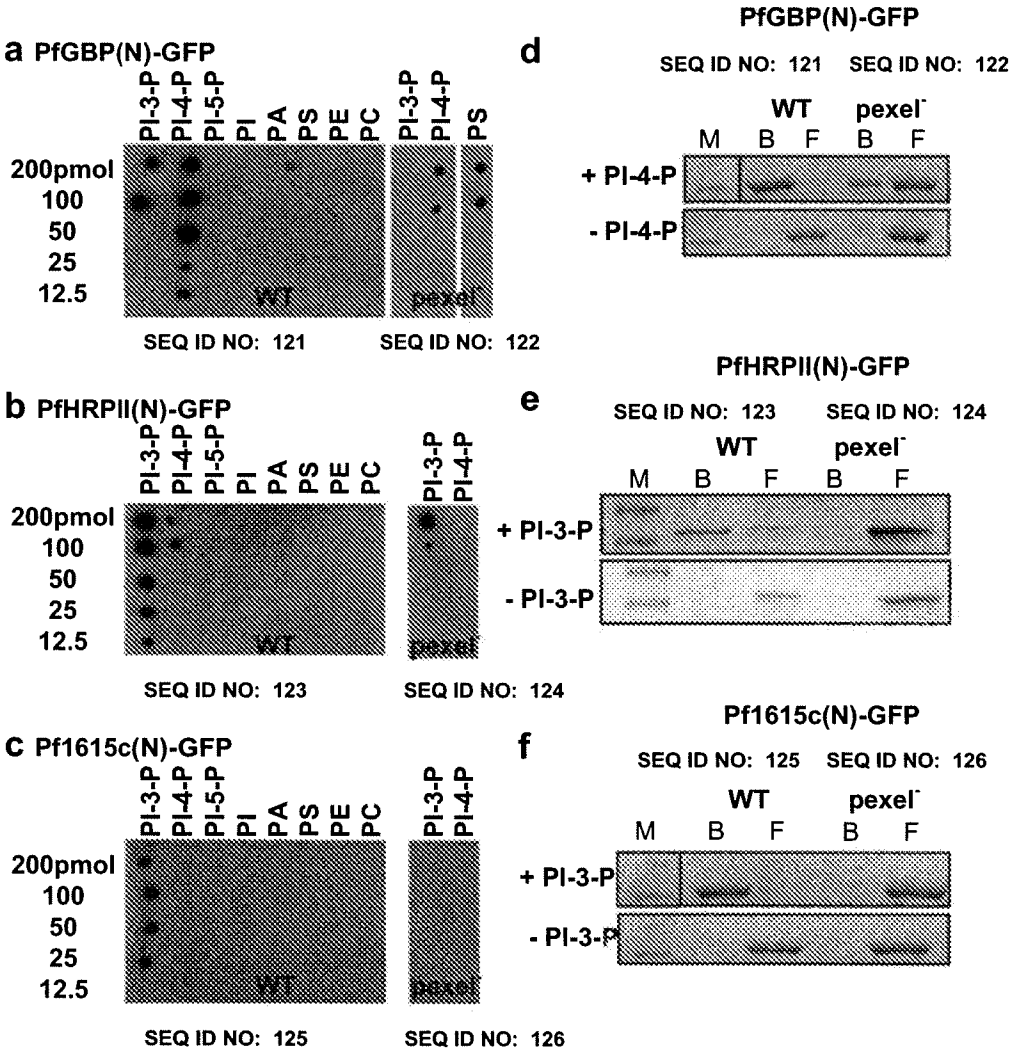
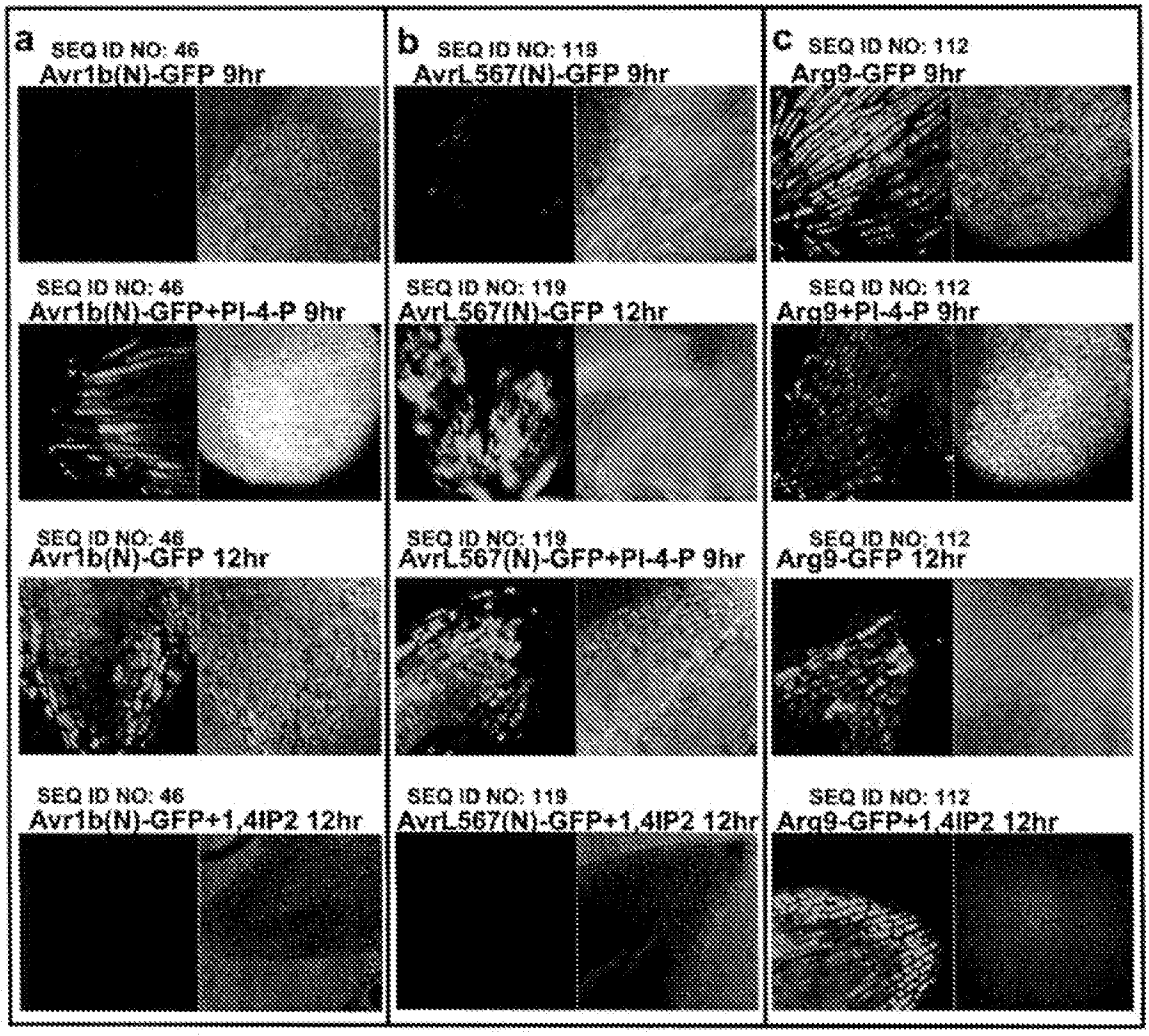

a PfGBP(N)-GFP
PI-3-P
PI-4-P
PI-5-P
PI
PA
PS
PE
PC
PI-3-P
PI-4-P
PS
200pmol
100
50
25
12.5
WT
pexel-
SEQ ID NO: 121
SEQ ID NO: 122
b PfHRPII(N)-GFP
SEQ ID NO: 123
SEQ ID NO: 124
c Pf1615c(N)-GFP
SEQ ID NO: 125
SEQ ID NO: 126
d
PfGBP(N)-GFP
SEQ ID NO: 121 SEQ ID NO: 122
WT
pexel-
M B F B F
+ PI-4-P
- PI-4-P
e
PfHRPII(N)-GFP
SEQ ID NO: 123 SEQ ID NO: 124
+ PI-3-P
- PI-3-P
f
Pf1615c(N)-GFP
SEQ ID NO: 125 SEQ ID NO: 126

Figure 8A-F


a
SEQ ID NO: 46
Avr1b(N)-GFP 9hr
b
SEQ ID NO: 119
AvrL567(N)-GFP 9hr
c
SEQ ID NO: 112
Arg9-GFP 9hr
SEQ ID NO: 46
Avr1b(N)-GFP+PI-4-P 9hr
SEQ ID NO: 119
AvrL567(N)-GFP 12hr
SEQ ID NO: 112
Arg9+PI-4-P 9hr
SEQ ID NO: 46
Avr1b(N)-GFP 12hr
SEQ ID NO: 119
AvrL567(N)-GFP+PI-4-P 9hr
SEQ ID NO: 112
Arg9-GFP 12hr
SEQ ID NO: 46
Avr1b(N)-GFP+1,4IP2 12hr
SEQ ID NO: 119
AvrL567(N)-GFP+1,4IP2 12hr
SEQ ID NO: 112
Arg9-GFP+1,4IP2 12hr

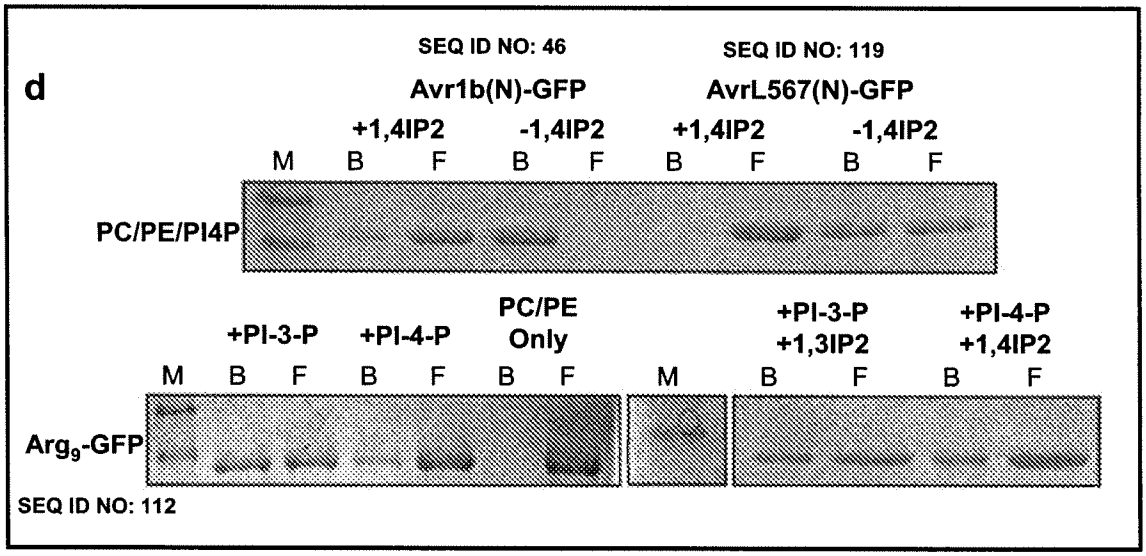

d
SEQ ID NO: 46
Avr1b(N)-GFP
SEQ ID NO: 119
AvrL567(N)-GFP
+1,4IP2
-1,4IP2
+1,4IP2
-1,4IP2
M B F B F B F B F
PC/PE/PI4P
+PI-3-P
+PI-4-P
PC/PE
Only
+PI-3-P
+1,3IP2
+PI-4-P
+1,4IP2
M B F B F B F M B F B F
Arg9-GFP
SEQ ID NO: 112

Figure 9 A-D

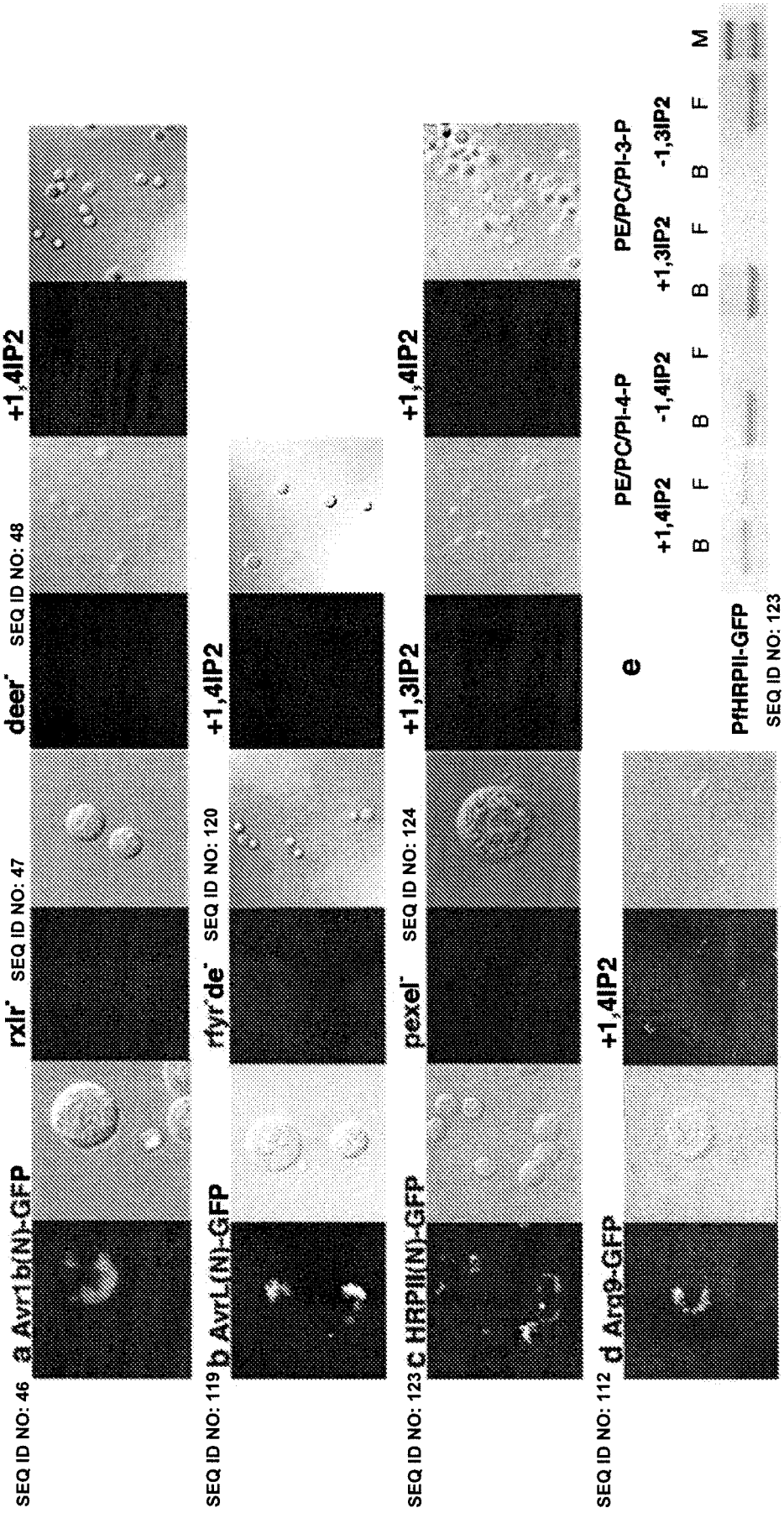

SEQ ID NO: 46
a Avr1b(N)-GFP
rxlr* SEQ ID NO: 47
deer* SEQ ID NO: 48
+1,4IP2
SEQ ID NO: 119
b AvrL(N)-GFP
rtyr*de* SEQ ID NO: 120
+1,4IP2
SEQ ID NO: 123
c HRPII(N)-GFP
pexel* SEQ ID NO: 124
+1,3IP2
+1,4IP2
SEQ ID NO: 112
d Arn9-GFP
+1,4IP2
e
PE/PC/PI-4-P
PE/PC/PI-3-P
+1,4IP2 -1,4IP2 +1,3IP2 -1,3IP2
B F B F B F B F M
PfHRPII-GFP
SEQ ID NO: 123

Figure 10 A-E

SEQ ID NO:	Plasmid Name	Sources	Construct	Construction strategy
1	pUN	pHAMT35N (*5*)	*Npt* II gene for G418 resistance fused to *P. sojae* rpL41 promoter in pUC19	Ham34 promoter of pHAMT35N replaced by *P. sojae* rpL41 promoter using primers UF and UR
2	pHamAvr1b	pHAMT35N	*Avr1b*-1 gene fused to Ham34 promoter and terminator for *P. sojae* transformation in pUC19	NptII gene of pHAMT35N replaced with *P. sojae* *Avr1b*-1 using PrimerC and PrimerD
3	M1	pHAMT35N	*Avr1b*-1(RXLR1^{AAAA}) mutant fused to Ham34 promoter and terminator	Same as pHamAvr1b with mutation introduced by primers motif1F and motif1R
4	M2	pHAMT35N	*Avr1b*-1(RXLR2^{AAAA}) mutant fused to Ham34 promoter and terminator	Same as pHamAvr1b with mutation introduced by primers motif2F and motif2R
5	M1+2	pHAMT35N	*Avr1b*-1(RXLR1^{AAA}A,2^{AAAA}) mutant fused to Ham34 promoter and terminator	Same as pHamAvr1b with mutation introduced by primers motif1+2F and motif1+2R
6	M3	pHAMT35N	*Avr1b*-1(dEER^{A6}) mutant fused to Ham34 promoter and terminator	Same as pHamAvr1b with mutation introduced by primers motif3F and motif3R
7	pHamAvh341	pHAMT35N	*HpAvh341* fused with C-terminal domain of *Avr1b*-1, driven by Ham34 promoter and terminator	NptII gene replaced with HpAvh341-Avr1b fusion using primers HpAvh341F, HpAvh341R and PrimerD
8	pHamAvh171	pHAMT35N	Ps *Avh171* (=*Avr4/6*) fused with C-terminal of *Avr1b*-1 driven by Ham34 promoter and terminator	NptII gene replaced with Ps Avh171-Avr1b fusion using primers PsAvh171F, PsAvh171R and PrimerD
9	pHamAvr1bCt	pHAMT35N	C-terminus of *Avr1b*-1 gene fused to Ham34 promoter and terminator for *P. sojae* transformation in pUC19	NptII gene of pHAMT35N replaced with *P. sojae* *Avr1b*-1 C terminus using primers 1bRXLR(-)F and PrimerD
10	UNM1	pUN and M1	Ham34::Avr1b (RXLR1^{AAAA}) inserted together with rpL41::NptII for *P. sojae* transformation	Selection and Gus expression cassettes in pCambia1305.2 replaced by inserts from pUN and M1, respectively
11	UNM2	pUN and M2	Ham34::Avr1b (RXLR2^{AAAA}) inserted together with rpL41::NptII for *P. sojae* transformation	Selection and Gus expression cassettes in pCambia1305.2 replaced by inserts from pUN and M2, respectively
12	UNM1+2	pUN and M1+2	Ham34::Avr1b (RXLR1^{AAAA}A,2^{AAAA}) inserted together with rpL41::NptII for *P. sojae* transformation	Selection and Gus expression cassettes in pCambia1305.2 replaced by inserts from pUN and M1+2, respectively
13	UNM3	pUN and M3	Ham34::Avr1b(dEER^{A6}) inserted together with rpL41::NptII for *P. sojae* transformation	Selection and Gus expression cassettes in pCambia1305.2 replaced by inserts from pUN and M3, respectively
14	pCambiaAvr1b	pCambia1305.2	CaMV 35S promoter fused to full length *Avr1b*-1 in GUS-containing vector	HMT gene in pCambia1305.2 replaced by *Avr1b*-1 using primers Avr1bfull_F and Avr1bR
15	pCambiaM2	pCambia1305.2	CaMV 35S promoter fused to *Avr1b*-1 (RXLR2^{AAAA}) in GUS-containing vector	HMT gene in pCambia1305.2 replaced by Avr1b(RXLR2) using primers Avr1bfull_F and Avr1bR
14	pCambia-mAvr1b	pCambia1305.2	CaMV 35S promoter fused to leaderless *Avr1b*-1 in GUS-containing vector	HMT gene in pCambia1305.2 replaced by mature *Avr1b*-1 using primers Avr1bF and Avr1bR
15	pCambia-mM2	pCambia1305.2	CaMV 35S promoter fused to leaderless *Avr1b*-1(RXLR2^{AAAA}) in GUS-containing vector	HMT gene in pCambia1305.2 replaced by mAvr1b(RXLR2) using primers Avr1bF and Avr1bR
16	pCa-GUS(-)	pCambia1305.2	Empty vector as the control	GUS expression cassette was removed by *Sph* I restriction and re-ligation for co-transformation experiments
17	pCaAvr1b	pCambiaAvr1b	CaMV 35S promoter fused to *Avr1b*-1 in GUS-free vector	
18	pCaM2	pCambiaM2	CaMV 35S promoter fused to *Avr1b*-1 (RXLR2^{AAAA}) in GUS-free vector	
19	pCa-mAvr1b	pCambia-mAvr1b	CaMV 35S promoter fused to leaderless *Avr1b*-1 in GUS-free vector	
20	pCa-mM2	pCambia-mM2	CaMV 35S promoter fused to leaderless *Avr1b*-1(RXLR2^{AAAA}) in GUS-free vector	
21	pUCAvr1b	pCambiaAvr1b	CaMV 35S promoter fused to *Avr1b*-1 with *Xma* I site	Avr1b expression cassette was inserted into pUC19 by PCR with primers Avr1b_EcoRI and Avr1b_HindIII
22	pUCAvr1bXK	pUCAvr1b	*Kpn* I site added to 3' end pUCAvr1b to facilitate later constructions	Amplification with primers Avr1b_EcoRI, Avr1b_HindIII and Avr1b_genegun_KpnI
23	pUCmAvr1b(dEER)	pUCAvr1bXK and M3	CaMV 35S promoter fused to mature *Avr1b*-1(dEER^{A6}) using *Xma* I and *Kpn* I sites	Ligation of *Xma* I-*Kpn* I fragment from M3 into pUCAvr1bXK

Figure 11A

24	pUCmAvr1bGFP	pAcGFP1-N	Fusion of leaderless *Avr1b*-1 with *AcGFP* from pAcGFP1-N (Clontech, # 632485) placed under control of CaMV 35S promoter	AcGFP fused to relevant Avr1b sequence by PCR with primers Avr1bF or Avr1bfull_F plus GFPF and GFPR. PCR product inserted into pUCAvr1bXK using *Xma* I and *Kpn* I
25	pUCsAvr1bGFP	pUCAvr1bXK	Fusion of normal *Avr1b*-1 with *AcGFP* placed under control of CaMV 35S promoter	
26	pUCGFP	pUCsAvr1bGFP	CaMV 35S promoter fused to *AcGFP*	Deletion of Avr1b sequences by self-ligation of *Xma I/Age* I restricted pUCsAvr1bGFP
27	pUCsGFP		CaMV 35S promoter fusion of Avr1b secretory leader to *AcGFP* placed under control of CaMV 35S promoter	Deletion of Avr1b sequences encoding mature protein by self-ligation of *Nco* I restricted pUCsAvr1bGFP
28	pUCM2GFP	pUCAvr1bXK	*Avr1b*-1(RXLR2^{AAAA})-*AcGFP* fusion placed under control of CaMV 35S promoter	AcGFP fused to relevant Avr1b sequence by PCR with primers Avr1bfull_F, GFPF and GFPR. PCR product inserted into pUCAvr1bXK using *Xma* I and *Kpn* I
29	pUCM3GFP		*Avr1b*-1(dEER^{A6})::*AcGFP* fusion placed under control of CaMV 35S promoter	
30	pUCAvr1b(R9)GFP		*Avr1b*-1(RXLR2^{Arg9})::*AcGFP* fusion placed under control of CaMV 35S promoter	
31	pUCAvr1b(TAT)GFP		*Avr1b*-1(RXLR2^{TAT})::*AcGFP* fusion placed under control of CaMV 35S promoter	
32	pUCNgo	pUCAvr1bXK	Addition of NgoMIV site upstream of RFLR in N-terminus of Avr1b (T150C) in pUCAvr1bXK to facilitate later manipulations	GCCGGT in was mutated to GCCGGc using primer Avr1bSac
33	pUCRFRL	pUCAvr1bXK	CaMV 35S promoter fused to *Avr1b*-1 genes carrying indicated mutations	Wild type Avr1b in pUCAv1bXK was replaced with the corresponding mutants using *Xma* I and *Kpn* I
34	pUCFRLR			
35	pUCRFAR			
36	pUCQFLQ			
37	pUCQFLR	pUCNgo		Avr1b in pUCNgo was replaced with the corresponding mutants by *NgoMIV*I and *Kpn* I
38	pUCKFLR			
39	pUCRFLQ			
40	pUCRFVR			
41	pUCAvr9R			
42	pUCAvrTAT			
43	pUCAvrPfGBP	pUCAvr1bXK		Wild type Avr1b in pUCAvr1bXK was replaced with the corresponding mutants using *Xma* I and *Kpn* I
44	pUCAvrPfHRP			
45	pUCAvrPf1615			
46	pAvr1bGFP	pR9GFP	Replacement of codons encoding nine arginine motif in pR9GFP with Avr1b residues 33 to 71 from wild type Avr1b or from Avr1b RXLR or dEER mutants	Amplification with primers AvrGFPF, AvrGFPR1 and AvrGFPR2
47	pAvr1b(M1+2)GFP			
48	pAvr1b(M3)GFP			

Figure 11B

SEQ ID NO:	Name	Applications	Sequence (from 5' to 3')*
49	UF	*Hind* III and *Xma* I sites added to rpL41 promoter to drive G418 resistance gene	ataagcttgaatTCTGGCGTTCATCTCCGACG
50	UR		ttcccgggTGGATGCTCAGATGctagcGTC
51	HamF	*Bremia* Ham34 promoter internal primer	TTCTCCTTTTCACTCTCACG
52	HamR	*Bremia* Ham34 Terminator internal primer	AGACACAAAATCTGCAACTTC
53	Avr1bReF	*Avr1b*-1 internal primers for PCR	ACCTTCAGCGTGACTGACCT
54	Avr1bReR		GCGATTGCCAACCAGTTCT
55	ActinF	*P. sojae Actin* gene internal primers for the reference in RT-PCR	CGACATCCGTAAGGACCTGT
56	ActinR		TTCGAGATCCACATCTGCTG
57	PrimerC	*Kpn* I sites added flanking *Avr1b*-1 for insertion downstream of HAM34 promoter	ggggtaccgacaacaATGCGTCTATCTTTTGTGCT
58	PrimerD		ggggtaccTCAGCTCTGATACCGGTGAA
59	Avr1bF	*Xho* I site and initiation codon added to 5' end of mature *Avr1b*-1 for insertion downstream of CaMV 35S promoter	atcgcactcgagctttcgcagatcccgggggcaatgagatatgACTGAGTACTCCGACGAA
60	Avr1bR	*Xho* I site to 3' end of *Avr1b*-1 for insertion downstream of CaMV 35S promoter	atcgcactcgagcttgtcgatcgacagatccggtcggcatctactTCAGCTCTGATACCGGTG
61	Avr1bfull_F	Same as Avr1bF but for secretory *Avr1b*-1	atcgcactcgagctttcgcagatcccgggggcaatgagatATGCGTCTATCTTTTGTG
62	Motif1F	Introduction of RXLR1^{AAAA} mutation, creating a *Pst* I site	TCGTCCGTgctgcagctgctAACGGCGACATTGCCGGTGG
63	Motif1R		TCGCCGTTagcagctgcagcACGGACGAGATCTGGAGATT
64	Motif2F	Introduction of RXLR2^{AAAA} mutation, creating a *Pst* I site	CCGGTGGAgctgcagctgctGCTCATGAAGAGGACGATGC
65	Motif2R		TCATGAGCagcagctgcagcTCCACCGGCAATGTCGCCGT
66	Motif1+2F	Introduction of RXLR1^{AAAA} & RXLR2^{AAAA} mutations, creating 2 *Pst* I sites	gctgcagctgctAACGGCGACATTGCCGGTGGAgctgcagctgctGCTCATGAAGAGGACGATGC
67	Motif1+2R		agcagctgcagcTCCACCGGCAATGTCGCCGTTagcagctgcagcACGGACGAGATCTGGAGATT
68	Motif3F	Introduction of dEER^{A6} mutation, creating a *Pst* I site	gctgcagcagctGCGGGGgctgctACCTTCAGCGTGACTGACCT
69	Motif3R		agcagcCCCCGCagctgctgcagcATGAGCTCGAAGAAATCTTC
70	HpAvh341F	5' flanking *Xma* I site added to Hp Avh341 for insertion downstream of HAM34 promoter	attcccggggacaacaATGCGACTCCACTACGTG
71	HpAvh341R	For fusion with *Avr1b*-1	GTCAGTCACGCTGAAGGT|ATCGAGAACGCCATGCCCAT
72	PsAvh171F	5' flanking *Xma* I site added to PsAvh171 for insertion downstream of HAM34 promoter	attcccggggacaacaATGGGCCTCCACAAGGGCT
73	PsAvh171R	For fusion with *Avr1b*-1	GTCAGTCACGCTGAAGGT|TAGGTGGTGTAGTCCGAC
74	1bRXLR(-)F	For deletion of N-terminus of *Avr1b*-1	aacccgggacaacaatgACCTTCAGCGTGACTGAC
75	Avr1b_EcoRI	Primers for making moving Avr1b expression cassette into pUC19 vector with *Xma* I and *Kpn* I sites flanking *Avr1b*-1 gene	GGAGgaaTTcGCTGGCTGGTGGCAGGAT
76	Avr1b_HindIII		GTATTGGCTAGAGaAGCTTGCCA
77	Avr1b_genegun_KpnI		AGAAACTCGAGCTTGTCGATCGACAGATCCGGTCGGCAggTACcTCAGCTCTGATAC
78	M2_F1	Replacement of RFLR with mutation encoding RFRL	GAAGATTTCgaCttGCTCATGAAG
79	M2_R1		CTTCATGAGCaaGtcGAAATCTTC
80	M2_F2	Replacement of RFLR with the mutation encoding FRLR	TGCCGGTGGAtttagaCTTCGAGCTC
81	M2_R2		GAGCTCGAAGtctaaaTCCACCGGCA
82	M2_F3	Replacement of RFLR with the mutation encoding RFAR	GAAGATTTgcaCGAGCTCAT
83	M2_R3		ATGAGCTCGtgcAAATCTTC

Figure 12 A

84	M2_F4	Replacement of **RFLR with the mutation** encoding QFLQ	**GGTGGAcagTTTCTTCaa**GCTCATGAAG
85	M2_R4		AGCttGAAGAAActgTCCACCGGCAATG
86	GFPF	For fusion with *Avr1b*-1 to make *Avr1b*-1-*AcGFP*	CTTTCACCGGTATCAGAGC|ggtaccgccacc|atgGTGAGCA AGGGCGCCGAG
87	GFPR	Addition of *Kpn* I after stop codon of GFP	AAGgtaCCtcaCTTGTACAGCTCATCCAT
88	Avr1bSac	Addition of NgoMIV site upstream of RFLR	TCATGAGCTCGAAGAAATCTTCCgCCGGCAATG
89	AvrRFLQ	Replacement of RFLR with the mutation encoding RFLQ	CATTGCCGGcGGAAGATTTCTTCaAGC
90	AvrQFLR	Replacement of RFLR with the mutation encoding QFLR	CATTGCCGGcGGAcaATTTC
91	AvrRFVR	Replacement of RFLR with the mutation encoding RFVR	CATTGCCGGcGGAAGATTTgTTC
92	AvrKFLR	Replacement of RFLR with the mutation encoding KFLR	CATTGCCGGcGGAAagTTTCTTC
93	Avr9R1	Replacement of RFLR with the sequence encoding the 9 arginine motif	cgtcgacgtcggcgacgcGCTCATGAAGAGGACGATG
94	Avr9R2		CATTGCCGGcGGAcgacggcgacgtcgacgtcggcgacg
95	AvrTAT1	Replacement of RFLR with the sequence encoding the TAT motif	taagaaacgccgtcagcgacgtcgaGCTCATGAAGAGGACGATG
96	AvfrTAT2		CATTGCCGGcGGAtatggacgtaagaaacgccgtcagc
97	GBPF1	Replacement of RFLR-dEER with the sequence from Pf GBP130	gaaggagaagactacactccggaaaagcaagcaaaag|ACCTTCAGCGT GACTGACC
98	GBPF2		ctcgtatactggcagagggcgaagatacctgcgcaaggaaggagaagactacact
99	GBPF3		gtatgagaaagcggtagattacggctttcgagagtctcgtatactggcagag
100	GBPR		ctaccgctttctcatacttatc|TGCGTTGCAGGTCACGAC
101	HRPF1	Replacement of RFLR-dEER with the sequence from Pf HRPII	tgtcgacgatgcgcaccatgcacaccatgttgcagat|ACCTTCAGCGTGA CTGACC
102	HRPF2		aacctcaacaagagactgttgcacgagacacaagcacatgtcgacgatgcgcacc
103	HRPF3		acaataacctgtgtagtaagaatgctaaaggcttgaacctcaacaagagactg
104	HRPR1		actacacaggttattgttaaa|TGCGTTGCAGGTCACGAC
105	1615F1	Replacement of RFLR-dEER with the sequence from Pf 1615c	cagtcaatgcattacaagaaatgttagatgatagtgtc|ACCTTCAGCGTGA CTGACC
106	1615F2		ctcaagcagttggagttcatcacattggaagagaagacagtcaatgcattacaag
107	1615F3		aagatcaactcgtcatctacttatacacacagtagaatactcaagcagttggagttc
108	1615R1		agatgacgagttgatcttgttgtaact|TGCGTTGCAGGTCACGAC
109	AvrGFPF	Replacement of codons encoding nine arginine motif in pR9GFP with Avr1b residues 33 to 71 from wild type Avr1b or from Avr1b RXLR or dEER mutants	gatctagatctGTGGAATCTCCAGATCTC
110	AvrGFPR1		GTCATATGGATAGCCGGACAT|GTCAGTCACGCTGAA GGT
111	AvrGFPR2		GATCCCATGGAGCCAGCATAGTCTGGGACGTCATATG GATAGCCGGA

Figure 12B

SEQ ID NO:	Plasmid	Recipient	Expression	*Description**	Cloning Strategy†
112	Arg9-GFP		*E. coli*	*9Arg fused to the N-terminus of GFP.*	Received from H.J.Lee (Chang, M., Chou, J. C. & Lee, H. J. Cellular internalization of fluorescent proteins via arginine-rich intracellular delivery peptide in plant cells. Plant Cell Physiol 46, 482-8 (2005)
46	Avr1b(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 33-71 of Avr1b fused to the N-terminus of GFP.*	Described in Figure 11
47	Avr1b(N)(rxlr$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 33-71 of Avr1b fused to the N-terminus of GFP. RSLR and RFLR both mutated to AAAA.*	Described in Figure 11
48	Avr1b(N)(deer$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 33-71 of Avr1b fused to the N-terminus of GFP. EEDDAGER mutated to AAAAAGAA.*	Described in Figure 11
113	Avh5(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 19-63 of Avh5 fused to the N-terminus of GFP.*	2-Step PCR of Avh5 using Avh5-ExpGFP-BglII and Avh5R2 then Avh5-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
114	Avh5(N)(rflr$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 19-63 of Avh5 fused to the N-terminus of GFP. RFLR mutated to AAAA.*	2-Step PCR of Avh5RXLR mutant from pGEX4T1 using Avh5-ExpGFP-BGIII and Avh5R2 then Avh5-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
115	Avh5(N)(deer$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 19-63 of Avh5 fused to the N-terminus of GFP. DTDIVYEPK mutated to ATAIVYAPA*	2-Step PCR of Avh5dEER mutant from pGEX4T1 using Avh5-ExpGFP-BGIII and Avh5R2 then Avh5-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
116	Avr1k(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 21-108 of Avr1k fused to the N-terminus of GFP.*	2-Step PCR of Avr1k using Avr1k-ExpGFP-BglII and Avr1kR2 then Avr1k-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
117	Avr1k(N)(rslr$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 21-108 of Avr1k fused to the N-terminus of GFP. RSLR mutated to AAAA*	2-Step PCR of Avr1kRXLR mutant from pGEX4T1 using Avr1k-ExpGFP-BglII and Avr1kR2 then Avr1k-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
118	Avr1k(N)(deer$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 21-108 of Avr1k fused to the N-terminus of GFP. DDDANVSIENR mutated to AAAANVSIANA*	2-Step PCR of Avr1kdEER mutant from pGEX4T1 using Avr1k-ExpGFP-BglII and Avr1kR2 then Avr1k-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.

Figure 13A

	Plasmid	Recipient	Expression	*Description**	Cloning Strategy†
119	AvrL567(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 24-67 of AvrL fused to the N-terminus of GFP*	2-Step PCR of AvrL567 from pUCmAvrL567 using AvrL-ExpGFP-BglII and AvrL5R2 then AvrL-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
120	AvrL567(N)(rfyr$^-$de$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 24-67 of AvrL fused to the N-terminus of GFP. RFYRSPTASVILSGLVKVKWDNE mutated to AAAASPTASVILSGLVKVKWANA*	2-Step PCR of AvrL567m from pUCmAvrL567(rfyr$^-$de$^-$) using AvrL-ExpGFP-BglII and AvrL5R2 then AvrL-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
121	PfGBP(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 70-107 of PfGBP fused to the N-terminus of GFP*	2-Step PCR of PfGBP using PfGBP-ExpGFP-BglII and PfGBPR2 then PfGBP-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
122	PfGBP(N)(pexel$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 70-107 of PfGBP fused to the N-terminus of GFP. RILAE mutated to AAAAA*	3-Step PCR of PFGBP mutating RILAE to AAAAA using PfGBPF1 & PfGBPR2, then PfGBPF2 and PfGBPR2, then PfGBPExp-BglII and Linker2. BglII and NcoI cloning.
123	PfHRPII(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 27-63 of PfHRPII fused to the N-terminus of GFP*	2-Step PCR of PfHRPII using PfHRPII-ExpGFP-BglII and PfHRPIIR2 then PfHRPII-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
124	PfHRPII(N)(pexel$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 27-63 of PfHRPII fused to the N-terminus of GFP. RLLHE mutated to AAAAA*	4-Step PCR of PfHRPII mutating RLLHE to AAAAA using PfHRPIIF1 and PfHRPIIR2, then PfHRPIIF2 and PfHRPIIR2, then PfHRPIIF3 and PfHRPIIR2, then PfHRPII-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
125	Pf1615c(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 25-65 Pf1615c fused to the N-terminus of GFP*	2-Step PCR of Pf1615c using Pf1615c-ExpGFP-BglII and Pf1615cR2 then Pf1615c-ExpGFP-BglII and LinkerR2. BglII and NcoI cloning.
126	Pf1615c(N)(pexel$^-$)-GFP	Arg9-GFP	*E. coli*	*Amino acids 25-65 Pf1615c fused to the N-terminus of GFP. RILKQ mutated to AAAAA*	3-Step PCR of Pf1615c mutating RILKQ to AAAAA using Pf1615cF1 and Pf1615cR2, then Pf1615cF2 and Pf1615cR2, then Pf1615c-ExpGFP and LinkerR2. BglII and NcoI cloning.

Figure 13 B

	Plasmid	Recipient	Expression	*Description**	Cloning Strategy†
127	GST-Avr1k	pGEX4T1	*E. coli*	*mature Avr1k fused to the C-terminus of GST*	PCR of Avr1k using Avr1kF-GST-BamHI and Avr1kR-GST-SalI. BamHI and SalI.
128	GST-Avr1k(rslr$^-$)	pGEX4T1	*E. coli*	*mature Avr1k fused to the C-terminus of GST. RSLR mutated to AAAA*	2-Step PCR of Avr1kRXLR using Avr1kF-GST-BamHI and Avr1kR2-RXLR, and Avr1kF2-RXLR and Avr1kR-GST-SalI. Products were fused using Avr1kF-GST-BamHI and Avr1kR-GST-SalI.
129	GST-Avh5	pGEX4T1	*E. coli*	*mature Avh5 fused to the C-terminus of GST*	PCR of Avh5 using Avh5F-GST-BamHI and Avh5R-GST-SalI. BamHI and SalI
130	GST-Avh5(rflr$^-$)	pGEX4T1	*E. coli*	*mature Avh5 fused to the C-terminus of GST.RFLR mutated to AAAA*	2-Step PCR of Avh5 using Avh5F-GST-BamHI and Avh5-RXLRF, and Avh5-RXLRR and Avh5R-GST-SalI. Products were fused using Avh5F-GST-BamHI and Avrh5R-GST-SalI.
131	GST-Avh5(deer$^-$)	pGEX4T1	*E. coli*	*mature Avh5 fused to the C-terminus of GST. DTDIVYEPK mutated to ATAIVYAPA*	2-Step PCR of Avh5 using Avh5F-GST-BamHI and Avh5-dEERF, and Avh5-dEER and Avh5R-GST-SalI. Products were fused using Avh5F-GST-BamHI and Avrh5R-GST-SalI.
132	pUCsAvr1k	pUCAvr1bXK	Transient soybean expression	*mature Avr1k fused to the secreted leader of Avr1b in the transient assay vector*	PCR of Avr1k using Avr1kF-NcoI and Avr1kR-KpnI. NcoI and KpnI.
133	pUCsAvr1k(rslr$^-$)	pUCAvr1bXK	Transient soybean expression	*mature Avr1k fused to the secreted leader of Avr1b in the transient assay vector. RSLR mutated to AAAA*	PCR of Avr1k(rflr$^-$) using Avr1kF-NcoI and Avr1kR-KpnI. NcoI and KpnI.
134	pUCsAvr1k(deer$^-$)	pUCAvr1bXK	Transient soybean expression	*mature Avr1k fused to the secreted leader of Avr1b in the transient assay vector. AAAANVSIANA*	PCR of Avr1k(deer$^-$) using Avr1kF-NcoI and Avr1kR-KpnI. NcoI and KpnI.
135	pUCmAvr1k	pUCAvr1bXK	Transient soybean expression	*mature Avr1k cloned into transient bombardment vector with Avr1b signal peptide*	PCR of Avr1k using Avr1kF-XmaI and Avr1kR-Kpn.I XmaI and KpnI.
	Plasmid	**Recipient**	**Expression**	***Description****	**Cloning Strategy†**
136	pUCmAvr1k(rslr$^-$)	pUCAvr1bXK	Transient soybean expression	*mature Avr1k cloned transient bombardment vector with Avr1b signal peptide. RSLR mutated to AAAA*	PCR of Avr1k(rflr$^-$) using Avr1kF-XmaI and Avr1kR-KpnI. XmaI and KpnI.
137	pUCmAvr1k(deer$^-$)	pUCAvr1bXK	Transient	*mature Avr1k cloned into transient*	PCR of Avr1k(deer$^-$) using Avr1kF-XmaI

Figure 13 C

			soybean expression	*bombardment vector with Avr1b signal peptide.*	and Avr1kR-KpnI. XmaI and KpnI.
138	pUCsAvrL567-Avr1b	pUCNgo	Transient soybean expression	*Amino acids 24-67 of AvrL replaced the RXLR region of sAvr1b in the transient assay vector*	4-Step PCR Avr1b. AvrLF1 and Avr1b-KpnI, then AvrLF2 Avr1b-KpnI, then AvrLF3 and Avr1b-KpnI, then AvrLF4 and Avr1b-KpnI.
139	pUCsAvrL567(rfyr⁻de⁻)-Avr1b	pUCNgo	Transient soybean expression	*Amino acids 24-67 of AvrL replaced the RXLR region of sAvr1b in the transient assay vector. RFYR-DE of AvrL muated to AAAA-AA*	2-Step PCR AvrL567-Avr1b. AvrLF and AvrLmR. AvrLmF and Avr1b-KpnI. Fuse the two pieces using AvrLF and Avr1b-KpnI.
140	pUCmAvrL567-Avr1b	pUCAvr1bXK	Transient soybean expression	*Amino acids 24-67 of AvrL replaced the RXLR region of mAvr1b in the transient assay vector*	PCR pUCsAvrL567. mAvrLF and Avr1b-KpnI of. XmaI and KpnI.
141	pUCmAvrL567(rfyr⁻de⁻)-Avr1b	pUCAvr1bXK	Transient soybean expression	*Amino acids 24-67 of AvrL replaced the RXLR region of mAvr1b in the transient assay vector. RFYR-DE of AvrL muated to AAAA-AA*	PCR pUCsAvrL567(rfyr-de-). mAvrLF and Avr1b-KpnI. XmaI and KpnI.
142	pUCsAvrM-Avr1b	pUCAvr1bXK	Transient soybean expression	*N-terminus of AvrM replaced the RXLR region of sAvr1b in the transient assay vector.*	3-Step PCR.AvrMF1 and AvrMR, AvrMF2 and AvrMR, AvrMF3 and Avr1b-KpnI. NcoI and KpnI.
143	pUCsAvrPita-Avr1b	pUCAvr1bXK	Transient soybean expression	*N-terminus AvrPita replaced the RXLR region of sAvr1b in the transient assay vector.*	4-Step PCR. AvrPita. AvrPitaF and AvrPitaR1, AvrPitaF and AvrPitaR2. Avr1b. AvrPitaF1mid and Avr1b-KpnI of Avr1b. AvrPitaF2mid and Avr1b-KpnI. Fusion using both products. AvrPitaF and Avr1b-KpnI. NcoI and KpnI.
	Plasmid	**Recipient**	**Expression**	***Description****	**Cloning Strategy†**
144	pUCsAvr1b(KRRAYIER)	pUCAvr1bXK	Transient soybean expression	*RFLR motif of Avr1b was replaced with KRRAYIER by PCR and fused to pUCAvr1bXK. XmaI/KpnI*	PCR of Avr1b using Avr1b(KRRAYIER)F and Avr1b-KpnI. NgoMIV and KpnI.
145	pUCsAvr1b(RAALK)	pUCAvr1bXK	Transient soybean expression	*RFLR motif of Avr1b was replaced with by RAALK PCR and fused to pUCAvr1bXK. XmaI/KpnI*	PCR of Avr1b using Avr1b(RAALK)F and Avr1b-KpnI. NgoMIV and KpnI.
146	pUCsAvr1b (RLFKLIFKTDSTDIQ)	pUCAvr1bXK	Transient soybean expression	*RFLR motif of Avr1b was replaced with RLFKLIFKTDSTDIQ by PCR and fused to pUCAvr1bXK. XmaI/KpnI*	3-Step PCR. PCR of Avr1b Avr1bFull-F and Avr1b(RLFKLIFKTDSTDIQ)R. PCR of Avr1b Avr1b(RLFKLIFKTDSTDIQ)F and Avr1b-KpnI. PCR fusion of products. Avr1bFull-F and Avr1b-KpnI. XmaI & KpnI.
38	pUCsAvr1b(KFLR)	pUCAvr1bNgo	Transient	*RFLR motif of Avr1b was replaced*	Described in Figure 11

Figure 13 D

			soybean expression	*with KFLR by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	
147	pUCsAvr1b(HFLR)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with HFLR by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	PCR of Avr1b using Avr1b(HFLR)F and Avr1b-KpnI. NgoMIV and KpnI.
37	pUCsAvr1b(QFLR)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with QFLR by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	Described in Figure 11
148	pUCsAvr1b(RFFR)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RFFR by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	PCR of Avr1b using Avr1b(RFFR)F and Avr1b-KpnI. NgoMIV and KpnI.
149	pUCsAvr1b(RFYR)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RFYR by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	PCR of Avr1b using Avr1b(RFYR)F and Avr1b-KpnI. NgoMIV and KpnI.
150	pUCsAvr1b(RFIR)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RFIR by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	PCR of Avr1b using Avr1b(RFIR)F and Avr1b-KpnI. NgoMIV and KpnI.
151	pUCsAvr1b(RFMR)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RFMR by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	PCR of Avr1b using Avr1b(RFMR)F and Avr1b-KpnI. NgoMIV and KpnI.
40	pUCsAvr1b(RFVR)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RFVR by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	Described in Figure 11
	Plasmid	**Recipient**	**Expression**	***Description****	**Cloning Strategy†**
35	pUCsAvr1b(RFAR)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RFAR by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	Described in Figure 11
39	pUCsAvr1b(RFLQ)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RFLQ by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	Described in Figure 11
152	pUCsAvr1b(RFLG)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RFLG by PCR and fused to pUCAvr1bXK. NgoMIV/KpnI*	PCR of Avr1b using Avr1b(RFLG)F and Avr1b-KpnI. NgoMIV and KpnI.
153	pUCsAvr1b(RMVR)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RMVR by PCR and fused to pUCAvr1bXK NgoMIV/KpnI*	PCR of Avr1b using Avr1b(RMVR)F and Avr1b-KpnI. NgoMIV and KpnI
154	pUCsAvr1b(RLGT)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RLGT by PCR and fused to pUCAvr1bXK NgoMIV/KpnI*	PCR of Avr1b using Avr1b(RLGT)F and Avr1b-KpnI. NgoMIV and KpnI

Figure 13E

155	pUCsAvr1b(RLTQ)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with RLTQ by PCR and fused to pUCAvr1bXK NgoMIV/KpnI*	PCR of Avr1b using Avr1b(RLTQ)F and Avr1b-KpnI. NgoMIV and KpnI
41	pUCAvr1b(Arg9)	pUCAvr1bNgo	Transient soybean expression	*RFLR motif of Avr1b was replaced with 9 Arg residues by PCR and fused to pUCAvr1bXK NgoMIV/KpnI*	Described in Figure 11
224	Pita(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 26-100 of Avr Pita fused to the N-terminus of GFP*	2-Step PCR using PitaForGFP-BglII and PitaR2 and LinkerR2. BglII and NcoI.
225	Pitam(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 26-100 of Avr Pita fused to the N-terminus of GFP. Mutation of putative RXLR motifs* KRRAY, RYSQ, RAALK, HAVK, RLFR to SASAS, AASA, ASSAA, ASAA, AASA.	2-Step PCR using PitaForGFP-BglII and PitaR2 and LinkerR2. BglII and NcoI.
226	Ps87(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 23-69 of Ps87 fused to the N-terminus of GFP.*	2 Step PCR using 87ForGFP-BglII and 87R2 and LinkerR2. BglII and NcoI
227	Ps87m(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 23-69 of Ps87 fused to the N-terminus of GFP. KRLTG mutated to AAAAA*	2 Step PCR using 87FForGFP-BglII and 87R2 and LinkerR2. BglII and NcoI
228	AvrM(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 54-151 of AvrM fused to the N-terminus of GFP.*	2 Step PCR using AvrMForGFP-BglII and AvrMR2 and LinkerR2. BglII and NcoI
229	AvrMm(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 54-151 of AvrM fused to the N-terminus of GFP.* HPIQ, KSGT, KDMK, KALK, KLGT, RGFLR, KFLK, KLST, KNVR mutated to AAAA,AAAA,AAAA,AAAA,AAAA,AAAA,AAAA,AAAA	2 Step PCR using AvrMForGFP-BglII and AvrMR2 and LinkerR2. BglII and NcoI
230	P123(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 23- 59 of AvrP123 were fused to the N-terminus of GFP.*	2 Step PCR using P123ForGFPBglII and P123R2 and LinkerR2. BglII and NcoI.
231	P123m(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 23- 59 of AvrP123 were fused to the N-terminus of GFP. RLTQ was mutated to AAAA*	2 Step PCR using P123ForBglII and P123R2 and LinkerR2. BglII and NcoI.
232	P4(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 29-61 of AvrP4 were fused to the N-terminus of GFP.*	2 Step PCR using P4ForGFPBglII and P4R2 and LinkerR2. BglII and NcoI.
233	P4m(N)-GFP	Arg9-GFP	*E. coli*	*Amino acids 29-61 of AvrP4 were fused to the N-terminus of GFP.RDIQ was mutated to AAAA.*	2 Step PCR using P4ForGFPBglII and P4R2 and LinkerR2. BglII and NcoI.

Figure 13 F

SEQ ID NO:	Primer Name	Function	Sequence (5' to 3')*
156	LinkerR2	Recreates the linker region between GFP and Arg9 for RXLR regions for all GFP expression constructs. NcoI	GCATC**CCATGG**AGCCAGCATAGTCTGGGACGT CATATGGATAGCCGGACATGG
157	Avr1k-GFPExp-BglII	Forward primer for Avr1k RXLR Region. BglII site for cloning into GFP expression vector	ATGC**AGATCT**CTCACTTGCGCCACCTCCGAGC
158	Avr1kR2	Reverse primer for Avr1k RXLR region. Adds nucleotides to facilitate space between GFP and RXLR	CATATGGATAGCCGGACATGGTGGATCTGTC GACTCTGTTTTCAATCG
159	Avr1kdEERR2	Reverse primer for Avr1k RXLR region dEER mutant. Adds nucleotides to facilitate space between GFP and RXLR	CATATGGATAGCCGGACATGGTGGATCTGTCG ACTGCGTTTGCAAT
160	Avh5-GFPExp-BglII	Forward primer for Avh5 RXLR Region. BglII site for cloning into GFP expression vector	AGTC**AGATCT**ACAAGAGTCCCCGACGA
161	Avh5R2	Reverse primer for Avh5 RXLR region. Adds nucleotides to facilitate space between GFP and RXLR	CATATGGATAGCCGGACATGGTGGATCTG TCGACTTGCCAGGATTATG
162	AvrL-GFPExp-BglII	Forward primer for AvrL567 RXLR Region. BglII site for cloning into GFP expression vector	ATCG**AGATCT**ATGGGAATGGAACATG
163	AvrLRevMod1	Reverse primer for AvrL567 RXLR region. Adds nucleotides to facilitate space between GFP and RXLR	GATAGCCGGACATGGTGGATCTGTCGACCCAC A GGTCAGTCACGC
164	AvrLRevMod2	Adds remaining nucelotides for linker sequence between GFP Expression vector and RXLR region	TCCCATGGAGCCAGCATAGTCTGGGACGTCAT A TGGATAGCCGGACATGGTGGATCT
165	PfGBP-GFPExp-BglII	Forward primer for PfGBP RXLR Region. BglII site for cloning into GFP expression vector	ATCG**AGATCT**GATAAGTATGAGAAAGCG
166	PfGBPR2	Reverse primer for PfGBP7 RXLR region. Adds nucleotides to facilitate space between GFP and RXLR	CATATGGATAGCCGGACATGGTGGATCTGTCG AC CGTTTGCTTGCTTTTCCGG
167	PfGFPm-GFPExp-F1	Forward primer 1 for PfGBP RXLR Region. RILAE to AAAAA	GATTACGGCTTTCGAGAGTCTgcagcagcagcagca GGCGAAGATACCTGCG
168	PfGFPm-GFPExp-F2	Forward primer 2 for PfGBP RXLR Region. RILAE to AAAAA	ATGCAGATCTGATAAGTATGAGAAAGCGGTAG ATT ACGGCTTTCGAGAGTCT
169	PfHRPII-GFPExp-BglII	Forward primer for PfHRPII RXLR Region. BglII site for cloning into GFP expression vector	ATGCAGATCTTTTAACAATAACC
170	PfHRPIIR2	Reverse primer for PfHRPII RXLR region. Adds nucleotides to facilitate space between GFP and RXLR	CATATGGATAGCCGGACATGGTGGATCTGTCG AC ATCTGCAACATGG
171	PfHRPIIm-GFPExp-	Forward primer 1 for PfGBP RXLR Region. RLLHE to	GGCTTGAACCTCAACAAGgcagcagcagcagca

Figure 14 A

	F1	AAAAA	ACACAAGCACATG
172	PfHRPIIm-GFPExp-F2	Forward primer 2 for PfGBP RXLR Region. RLLHE to AAAAA	GTAGTAAGAATGCTAAAGGCTTGAACCTCAACAAG
	Primer Name	**Function**	**Sequence (5' to 3')**
173	PfHRPIIm-GFPExp-F3	Forward primer 3 for PfGBP RXLR Region. RLLHE to AAAAA	ATGC**AGATCT**TTTAACAATAACCTGTGTAGTAAGA ATGCTAAAGGC
SEQ ID NO:	**Primer Name**	**Function**	**Sequence (5' to 3')**
174	Pf1615cGFPExp-BglII	Forward primer for Pf1615c RXLR Region. BglII site for cloning into GFP expression vector	ATGC**AGATCT**AGTTACAACAAGATCAAC
175	Pf1615cR2	Reverse primer for Pf1615c RXLR region. Adds nucleotides to facilitate space between GFP and RXLR	CATATGGATAGCCGGACATGGTGGATCTGTCGAC ACTATCATCTAAC
176	Pf1615cm-GFPExp-F1	Forward primer 1 for Pf1615c RXLR Region. RILKQ to AAAAA	CATCTACTTATACACACAGTgcagcagcagcagca TTGGAGTTCATCACATTGG
177	Pf1615cm-GFPExp-F2	Forward primer 2 for Pf1615c RXLR Region. RILKQ to AAAAA	ATGCAGATCTAGTTACAACAAGATCAACTCGTCATC TACTTATACACACAGTgcagca
178	Avr1kF-GST-BamHI	Forward primer for mature Avr1k. BamHI site for cloning into PGEX4T1 vector	GCGT**GGATCC**CTCACTTGCGCCACCTCCG
179	Avr1kR-GST-SalI	Reverse primer for mature Avr1k. SalI site for cloning into PGEX4T1 vector	TCGA**GTCGAC**TCAGATAATCATGATGCTGT
180	Avh5F-GST-BamHI	Forward primer for mature Avh5. BamHI site for cloning into PGEX4T1 vector	GCGT**GGATCC**ACAAGAGTCCCCGACGA
181	Avh5R-GST-SalI	Reverse primer for mature Avh5. SalI site for cloning into PGEX4T1 vector	TCGA**GTCGAC**TCACTTGGCCTCTTGCGAT
182	Avr1kF2-RXLR	Forward primer for producing RXLR mutation in Avr1k. RSLR to AAAA	CAgcggcggcggcgTCTCAAGCTACGAACG
183	Avr1kR2-RXLR	Reverse primer for producing RXLR mutation in Avr1k. RSLR to AAAA	GTTCGTAGCTTGAGAcgccgccgccgcTGTCGAGGTTGCCATGTC
184	Avr1kF2-dEER	Forward primer for producing dEER mutation in Avr1k. DDDANVSIENR to AAAANVSIANA	GGCTAACGTTTCGATTgcgAACgcgGGATGAACCCTTCAG
185	Avr1kR2-dEER	Reverse primer for producing dEER mutation in Avr1k. DDDANVSIENR to AAAANVSIANA	CGCAATCGAAACGTTAGCcgccgccgcCACGTTCGTAG
186	Avr1kF-NcoI	Forward primer for mature Avr1k. NcoI site for cloning into transient assay vector with Avr1b signal peptide	GCGT**CCATGG**GAGCTCACTTGCGCCACCTCCG
187	Avr1kR-KpnI	Reverse primer for mature Avr1k. KpnI site for cloning into transient assay vector	TCGA**GGTACC**TCAGATAATCATGATGCTGT

Figure 14 B

188	Avr1kF-XmaI	Forward primer for mature Avr1k. XmaI site for cloning into transient assay vector	AGAT**CCCGGG**GGGCAGTGAGATATGCTCACTT GCGCCACCTCCGAGCAACAG
189	AvrL567-F4	Forward primer for AvrL567 RXLR region NcoI site for cloning into transient assay vector with Avr1b signal peptide	ATCG**CCATGG**GAATGGAACATGTACCAGCAGA GTT GACCAGAGTCAGCGAAGGG
190	AvrL567-F3	Forward primer for AvrL567 RXLR region	ACCAGAGTCAGCGAAGGGTATACACGATTTTA CCGGTCCCCAACGGCTAGTGTAATAC
191	AvrL567-F2	Forward primer for AvrL567 RXLR region	CAACGGCTAGTGTAATACTGTCAGGATTGGTA AAGGTTAAATGGGATAATGAACAAATG
	Primer Name	**Function**	**Sequence (5' to 3')**
192	AvrL567-F1	Forward primer for AvrL567 RXLR region	TGGGATAATGAACAAATGACGATGCCGACCTT CAGCGTGACTGACCTGTGG
193	AvrL567-F	Forward primer for AvrL567 RXLR region XmaI site for cloning into transient assay vector	AGAT**CCCGGG**GGGCAATGAGATATGATGGAAC ATGTACCAGCAG
194	AvrL567MR	Reverse primer for AvrL567 mutation. RFYR-DxE to AAAA-AxA	TCCTGACAGTATTACACTAGCCGTTGG GGAcgcagcggctgcTGTATACCCTTCGCTG
195	AvrL567MF	Forward primer for AvrL567 mutation. RFYR-DxE to AAAA-AxA	GTGTAATACTGTCAGGATTGGTAAAGGTT AAATGGgctAATgcaCAAATGACGATGCCG
77	Avr1b-KpnI Dou et al[8]	Reverse Primer for Avr1b in the transient assay vector	AGAAACTCGAGCTTGTCGATCGACAGATCCGG TCGGCA**ggTACc**TCAGCTCTGATAC
196	Avr1bFull-F Dou et al[8]	Forward primer for Avr1b with signal peptide. XmaI site	atcgcactcgagctttcgcagat**cccggg**gggcaatgagatATG CGTCTATCTTTTGTG
197	AvrPitaF	Forward primer for fusing N-terminus of AvrPita wild type into Avr1b	ATCG**CCATGG**GACACCCAGTTTACGATTACAAT C
198	AvrPitaF2mid	Forward primer for fusing N-terminus of AvrPita wild type into Avr1b	CACGCCGTTAAAAATGACAATCGGTTATTTAGA T TAATCTTTAAAACTGACAGCACAG
199	AvrPitaF1mid	Forward primer for fusing N-terminus of AvrPita wild type into Avr1b	CTTTAAAACTGACAGCACAGATATTCAAAACAC C TTCAGCGTGACTGACCTG
200	AvrPitaR1	Reverse primer for fusing N-terminus of AvrPita wild type into Avr1b	CCACGAGGCGAGCTCGGCACAACTTTTTAGCG C GGCACGAATTTC
201	AvrPitaR2	Reverse primer for fusing N-terminus of AvrPita wild type into Avr1b	CCGATTGTCATTTTTAACGGCGTGATAGCCCCA C GAGGCGAGCTCGGCAC
202	AvrM-F3	Forward primer for fusing N-terminus of AvrM wild type into Avr1b	ATCG**CCATGG**GAAACAACCTTGGAACAGTACC GGA

Figure 14 C

			TGTGCCACATC
203	AvrM-F2	Forward primer for fusing N-terminus of AvrM wild type into Avr1b	CAGTACCGGATGTGCCACATCAAATTCCAAATG A CAAAAGTGGTACTCCTGCCATTG
204	AvrM-F1	Forward primer for fusing N-terminus of AvrM wild type into Avr1b	CAAAAGTGGTACTCCTGCCATTGAAGACCCAA AA GATATGAAAGGATTCAATAAGGCTC
205	AvrM-R	Reverse Prmer for fusing N-terminus of AvrM wild type into Avr1b	CCACAGGTCAGTCACGCTGAAGGTTCGTACAT TT TTATCTTTTTCTGCCATGTATG
206	Avr1b(KRRAYIER) F	Forward primer for replacing RFLR motif of Avr1b with KRRAYIER	GACATT**GCCGGC**GGAaaaaggcgggcttatatt gaacgcGCTCATGAAGAGGACGATGC
207	Avr1b(RAALK)F	Forward primer for replacing RFLR motif of Avr1b with RAALK	GACATT**GCCGGC**GGAcgtgccgcgctaaaaAG TGCTCATGAAGAGGACGATGC
208	Avr1b (RLFKLIFKTDSTDI Q)F	Forward primer for replacing RFLR motif of Avr1b with RFLKLIFKTDSTDIQ	ctttaaaactgacagcacagatattcaaGCTCATGA AGAGGACGATGCG
	Primer Name	**Function**	**Sequence (5' to 3')**
209	Avr1b (RLFKLIFKTDSTDI Q)R	Reverse primer for replacing RFLR motif of Avr1b with RFLKLIFKTDSTDIQ	ctgtgctgtcagttttaaagattaatttaaataaccgTCC **GCCGGC**AATGTCGCCGTTC
210	Avr1b(KFLR)F Dou et al[8]	Forward primer for replacing RFLR motif of Avr1b with KFLR	CATT**GCCGGc**GGAAagTTTCTTC
211	Avr1b(HFLR)F	Forward primer for replacing RFLR motif of Avr1b with HFLR	CGACATT**GCCGGC**GGAcacTTTCTTCGAGCTCA TGAAGAGGACGATG
212	Avr1b(QFLR)F Dou et al [8]	Forward primer for replacing RFLR motif of Avr1b with QFLR	CATT**GCCGGc**GGAcaATTTC
213	Avr1b(RFFR)F	Forward primer for replacing RFLR motif of Avr1b with RFFR	CGACATT**GCCGGC**GGAAGATTTtTTCGAGCTCA TGAAGAGGACGATG
214	Avr1b(RFYR)F	Forward primer for replacing RFLR motif of Avr1b with RFYR	CGACATT**GCCGGC**GGAAGATTTtaTCGAGCTCA TGA AGAGGACGATG
215	Avr1b(RFIR)F	Forward primer for replacing RFLR motif of Avr1b with RFIR	CGACATT**GCCGGC**GGAAGATTTaTTCGAGCTC ATG AAGAGGACGATG
216	Avr1b(RFMR)F	Forward primer for replacing RFLR motif of Avr1b with RFMR	CGACATT**GCCGGC**GGAAGATTTaTgCGAGCTC ATG AAGAGGACGATG
217	Avr1b(RFVR)F	Forward primer for replacing RFLR motif of Avr1b with	CATT**GCCGGc**GGAAGATTTgTTC

Figure 14D

	Dou et al [8]	RFVR	
218	Avr1b(RFAR)F Dou et al [8]	Forward primer for replacing RFLR motif of Avr1b with RFAR	CATT**GCCGGc**GGAAGATTTgcTCG
219	Avr1b(RFLQ)F Dou et al [8]	Forward primer for replacing RFLR motif of Avr1b with RFLQ	CATT**GCCGGc**GGAAGATTTCTTCaAGC
220	Avr1b(RFLG)F	Forward primer for replacing RFLR motif of Avr1b with RFLG	CGACATT**GCCGGC**GGAAGATTTCTTgGAGCTC AT GAAGAGGACGATG
221	Avr1b(RMVR)F	Forward primer for replacing RFLR motif of Avr1b with RMVR	CGACATT**GCCGGC**GGAAGAaTggTTCGAGCTCA TGAAGAGGACGATG
222	Avr1b(RLGT)F	Forward primer for replacing RFLR motif of Avr1b with RLGT	CGACATT**GCCGGC**GGAAGAcTTggaacAGCTCA TGAAGAGGACGATG
223	Avr1b(RLTQ)F	Forward primer for replacing RFLR motif of Avr1b with RLTQ	CGACATT**GCCGGC**GGAAGAcTTacTCaAGCTCA TGAAGAGGACGATG
234	P4ForGFPExp-BglII	Forward primer to PCR AvrP4 or AvrP4 mutant	atgcagatctGAATTCTTAGAGGATGCACGAGATATCC
235	P4RevGFPR2	Reverse primer to PCR AvrP4 or AvrP4 mutant	CTGACAGCTCGCGCGATAGGCAAGAAAAAGTCGACAGA TCCACCATGTCCGGCTATCCATATG
236	P123ForGFPExp-BglII	Forward primer to PCR AvrP123 or AvrP123 mutant	atgcagatctCAGTATGTTGTTGATCCAGG
237	P123RevGFPR2	Reverse primer to PCR AvrP123 or AvrP123 mutant	CATATGGATAGCCGGACATGGTGGATCTGTCGACACAA CTTGGGGTGGCTTCACATTCCAC
238	Ps87ForGFP-BglII	Forward primer to PCR Ps87 or Ps87 mutant	ATGCAGATCTACTCTGTTGAACGTGACT
239	Ps87RevGFP2	Reverse primer to PCR Ps87 or Ps87 mutant	CATATGGATAGCCGGACATGGTGGATCTGTCG ACGTTTTGAATATCTGTGC
240	AvrMForGFP-BglII	Forward primer to PCR AvrM or AvrM mutant	atgcagatctATGGGAAACAACCTTGGAACAGTACCG
241	AvrMRevGFPR2	Reverse primer to PCR AvrM or AvrM mutant	CATATGGATAGCCGGACATGGTGGATCTGTCGACTCGT ACATTTTTATCTTTTTCTGCCATG
242	AvrPitaForGFP-BglII	Forward primer to PCR AvrPita or AvrPita mutant	AGTCAGATCTCACCCAGTTTACGATTACAATCC
243	AvrPitaRevGFPR2	Forward primer to PCR AvrPita or AvrPita mutant	CATATGGATAGCCGGACATGGTGGATCTGTCG ACGTTTTGAATATCTGTGC
244	Ps87m8R	For replacing KRLTG with AAAAA	GAAACgcagcagctgcggcaAAACCTCGAGGGCAAT GTGTC
245	Ps87m8F	For replacing KRLTG with AAAAA	GTTTtgcagccgctgctgcGTTTCTCTCTTTCATCATT TC
246	AvrMmR1	Reverse Primers for replacing HPIQ, KDMK, KALK, KLGT, RGFLR, KFLK, KLST motifs in AvrM with	CATTtgcagatgaagcTGGCACATCCGGTACTGTTC

Figure 14E

247	AvrMmR2	ASSA, SASA, SAAA, AAAA, SAASA, SAAS, ASSA	cagcgctATTGAATCCtgctgaggcactTGGGTCTTCAA TGGCAGGAG
248	AvrMmR3		GACGAagccgcagcggcTTCGGATTCTGGAGTAGAt g
249	AvrMmR4		CAAAAGGcgcagaggctgcactGTCAAATTCTGGTTG AGGGATC
250	AvrMmR5		gatgcCTGAACTTGGTCCGGactagcggctgaCATTTT TGCTCCAAAAGGcgcag
251	AvrMmF1	Reverse Primers for replacing HPIQ, KDMK, KALK, KLGT, RGFLR, KFLK, KLST motifs in AvrM with ASSA, SASA, SAAA, AAAA, SAASA, SAAS, ASSA	GTGCCAgcttcatctgcaAATGACAAAAGTGGTACTC CTG
252	AvrMmF2		cagcaGGATTCAATagcgctgccgcaTCTACTCCAGA ATCCGAAAAAC
253	AvrMmF3		GAAgccgctgcggctTCGTCAGTTGAAGGGATCCCT C
254	AvrMmF4		Cagtgcagcctctgcg CCTTTTGGAGCAAAAATG
255	AvrMmF5		gtCCGGACCAAGTTCAGgcatcttcagccGATGATCT CATCACATACATG
256	AvrP4mR	Reverse primer used to replace RDIQ in AvrP4 with AAAA	
257	AvrP4mF	Forward primer used to replace RDIQ in AvrP4 with AAAA	
258	AvP123mF	Forward primer used to replace RLTQ in AvrP123 with AAAA	
259	AvrP123mR	Reverse primer used to replace RLTQ in AvrP123 with AAAA	

Figure 14F

SEQ ID NO:	Protein		RXLR Motifs of each fungal or mutations therein	PA Binding	Root Cell Entry
224	AvrPita-GFP	wild-type	KRRAY, RYSQ, RAALK, HAVK, RLFR	Y	Y
225	AvrPitam-GFP	mutant	SASAS, AASA, ASSAA, ASAA, AASA	N	N
226	Ps87-GFP	wild-type	KRLTG	Y	NT
227	Ps87m-GFP	mutant	AAAAA	N	NT
228	AvrM-GFP	wild-type	HPIQ, KSGT, KDMK, KALK, KLGT, RGFLR, KFLK, KLST, KNVR	Y	NT
229	AvrMm-GFP	mutant	AAAA,AAAA,AAAA,AAAA,AAAA,AAAAA,AAAA,AAAA	N	NT
230	AvrP123-GFP	wild-type	RLTQ	Y	NT
231	AvrP123m-GFP	mutant	AAAA	N	NT
232	AvrP4-GFP	wild-type	RDIQ	Y	NT
233	AvrP4m-GFP	mutant	AAAA	N	NT

Figure 15