The Molecular Basis for Recognition of Oomycete Effectors in *Arabidopsis*

by

Ksenia V. Krasileva

The dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Microbiology

and the Designated Emphasis

in

Computational Biology

in the

Graduate Division

of the

University of California, Berkeley

~

Committee in charge:

Professor Brian J. Staskawicz, Chair

Professor Kimmen Sjolander

Professor Greg Barton

Spring 2011

ABSTRACT

"The Molecular Basis for Recognition of Oomycete Effectors in *Arabidopsis"*

by

Ksenia V. Krasileva, Doctor of Philosophy in Microbiology

University of California, Berkeley

Professor Brian J. Staskawicz, Chair

This Dissertation presents a comprehensive study of how oomycete pathogens are recognized by the plant model organism *Arabidopsis thaliana*. The key molecular factors in this study include pathogen derived effector molecules, and plant disease resistance genes; specifically, the effector, Arabidopsis thaliana recognized 1 (ATR1) and cognate plant resistance gene, Recognition of *Peronospora parasitica* 1 (RPP1). Presented data, gathered using genetics, biochemistry, structural biology, natural history and human logic, provide insights into molecular basis for recognition of ATR1 by RPP1. It is demonstrated that the in planta physical association between ATR1 and RPP1 leads to activation of plant immunity. The main regions of RPP1 responsible for effector recognition and downstream activation of immunity are elucidated. Examination of natural variation among ATR1 and RPP1 variants and subsequent heuristic mutational analyses yielded proteins with reversed activity, allowing manipulating the plant to expand its range of pathogen recognition. The 3-D structure of ATR1 was determined and showed that several independent regions could be recognized by RPP1 variants, illustrating adaptive components in plant immunity. Finally, the observed interactions between 83 different *Arabidopsis* sub-species and 5 strains of its native obligate oomycete pathogen provide show multi-level regulations of host / pathogen interactions, including developmental control of resistance in the host, and the ability of pathogen to suppress recognition of effectors and plant immunity.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I want to acknowledge the people who guided me in this journey through graduate school towards becoming a doctor in philosophy and a good scientist. I am grateful to Brian Staskawicz for being a fantastic PhD advisor and a very cheerful person, for always believing in me and allowing vast intellectual freedom, while being strict about experimental controls. I am deeply grateful to Doug Dahlbeck, a master in scientific technique and in a zen-like patience. My special acknowledgements go to Sophien Kamoun for introducing me to the big world of oomycetes and for setting an example of a supreme scientist. I am grateful to all fellow current and past members of the Staskawicz lab, especially to Lauriebeth Leonelli for being on the same track as me, thus being able to understand all the ups and downs of this journey, and to Gitta Coaker for introducing me to biochemistry. I want to thank my thesis committee, Kimmen Sjolander and Greg Barton, for giving great guidance, advice and encouragement. I am grateful to all of my collaborators and colleagues, who worked with me on several projects and added so much breadth to my studies. Finally, I want to heartily thank my family for never questioning my career choices and taking pride in everything that I do and my fiancé Daniil Prigozhin for always being by my side listening, formulating, coaching, encouraging, criticizing, advising, editing and cooking great meals.

INTRODUCTION

INTRODUCTION

Life does not proceed in complete isolation; all organisms cope to co-exist with each other, engaging in neutral, symbiotic and antagonistic relationships. An example of the antagonistic relationship is the host / pathogen interaction. It is marked by ongoing evolution of sophisticated virulence components in the pathogen and an immune system on the side of the host. Accumulating experimental data suggests that the fundamental paradigm underlying immunity is shared among organisms from different kingdoms [1]. Protective systems in prokaryotic or eukaryotic cells are based on the same principles of recognition of the foreign pathogen-born molecules and activation of defense responses. The first level of immunity in both plant and animal organisms relies on recognition of conserved pathogen-derived molecules (PAMPs), such as bacterial flagellin, viral RNA or DNA or fungal polysaccharides [2]. Although the receptors recognizing PAMPs in plants and animals appear to have evolved independently, the recognition is mediated by structurally similar leucine-rich repeat domains connected to variable signaling domains. Ability of pathogens to escape PAMP-based immunity has led to evolution of the more flexible receptors in the host, such as antibody-based adaptive immunity in animals. In plants, ability to detect diverse and rapidly evolving pathogenic molecules depends on a class of proteins termed Resistance proteins (R-proteins) [3, 4].

The existence of plant immunity that is based on genetically flexible components and ability to recognize variable pathogen-derived molecules was first postulated by Harold Henry Flor in the 1940s [5]. He developed gene-for-gene concept to explain interactions between flax, *Linum usitatissimum*, and flax rust, *Melampsora lini* [6]. Gene-for-gene model states that induction of defense responses occurs only if both the plant and the pathogen have necessary genetic components. Thus, a pathogen can escape plant immunity by modifying the molecules it exposes to the host, and a plant needs to catch up by genetic modifications of its receptors, R-genes. This dynamic equilibrium follows The Red Queen Hypothesis, where both the parasite and the host have to evolve as fast as they can just to keep the relationships steady [7, 8]. This idea forms the theoretical foundation for the initial step in plant immunity, pathogen recognition.

The molecular mechanisms controlling this evolutionary dynamic system of pathogen recognition are elusive. Cloning of R-genes demonstrated that they share a common domain architecture of NBS-LRR proteins: a carboxyl-terminal Leucine Rich Repeat domain (LRR) is linked to a central Nucleotide Binding Region (NBS) and an amino-terminal Toll/Interleukin 1 homology domain or a coiled coil domain [9]. It has been also shown that diverse pathogen-derived virulence factors, recognized by R-proteins are commonly secreted and translocated inside the host and collectively called effector molecules. Individual studies of effector recognition in plants have not yet led to a unifying theory explaining the roles of different domains in R-proteins and mechanisms underlying co-evolution of effector recognition and pathogen escape.

The plant model organism *Arabidopsis thaliana* and its native obligate downy mildew pathogen *Hyaloperonospora arabidopsidis* provide an excellent model system for investigating gene-for-gene interactions [10]. Availability of *A. thaliana* sub-species, called ecotypes, is matched by genetically diverse strains of *H. arabidopsidis*. The ecotype-by-strain interactions are very similar to those observed by Flor in flax/flax rust system. Genome sequences of both *Arabidopsis* and *H. arabidopsidis* are available [11, 12]. Genetic analyses led to identification of several dozens of R-genes that confer resistance to *H. arabidopsidis* and are called RPP genes (for Recognition of *Peronospora parasitica*, former name of *H. arabidopsidis*) [13]. Genetic analyses in the pathogen are more complex due to its obligate biotrophic nature, and so far identified three effector molecules, ATR1, ATR5 and ATR13 (ATR stands for *Arabidopsis thaliana* recognized) [14]. However, bioinformatic analyses predict at least 130-150 ATR effector genes in *H. arabidopsidis* genome [11, 15], and their experimental verification is on the way. Investigating molecular mechanisms of ATR/RPP interactions will provide insights into host/pathogen interactions that can translate into better understanding of crop species and their parasites, as well as of the common principles of plant immunity.

During my dissertation years, I have primarily focused on interactions between the *H. arabidopsidis* effector molecule ATR1 and its cognate *Arabidopsis* R-protein, RPP1. The First Chapter presents the analyses of ATR1/RPP1 interactions on genetic and molecular biology level. It describes adoption of a surrogate transient expression system that allowed me to perform genetic analysis on ATR1, commonly unavailable in the pathogen due to its obligate nature. As a result of this, I was able to use biochemistry to determine the basis of ATR1 recognition by RPP1 [16]. This system also allowed defining the roles of different domains of RPP1 in effector recognition and subsequent downstream activation of defense responses [16]. The Second Chapter presents in depth analysis of the ATR1 recognition by RPP1. It shows a 3-D x-ray structure of ATR1, obtained in collaboration with the Tom Alber group at Berkeley, that has striking structural similarity to another effector, Avr3a11 from the potato blight pathogen *Phytophthora capsici*, which was solved by Sophien Kamoun and Mark Banfiled in the United Kingdom. The structure of ATR1 in combination with the genetic analyses based on natural variation between recognized and non-recognized ATR1, suggests that its recognition could have evolved independently in different RPP1 variants, since it is achieved through non-overlapping ATR1 surface epitopes. In the same Chapter, I present my structural model of RPP1 and show the first evidence for being able to control and expand its range of pathogen recognition through a controlled heuristics-based mutagenesis. Finally, the Third Chapter goes back to the interactions between different *Arabidopsis* ecotypes and pathogen strains. With indispensible help from my undergraduate research assistant, Connie Zheng, I have made a systematic analysis of interactions between 83 *Arabidopsis* ecotypes, collected from world-wide locations, and 5 strains of *H. arabidopsidis*. The analyses included investigation of juvenile and adult resistance in *Arabidopsis*, identification and quantification of different phenotypic outcomes of host-pathogen interactions, and the prevalence of ATR1 recognition. Together, these chapters present approaches, different in scale and technique, which have been successfully applied to investigate host - pathogen interactions.

CHAPTER I: Activation of an *Arabidopsis* resistance protein RPP1 is specified by the *in planta* association its LRR domain with the cognate oomycete effector ATR1.

Abstract

Activation of plant immunity relies on recognition of pathogen effectors by several classes of plant resistance proteins. To discover the underlying molecular mechanisms of effector recognition by the (TIR-NBS-LRR) RPP1 resistance protein, we adopted an *Agrobacterium*-mediated transient protein expression system in *Nicotiana tabacum*, which allowed us to perform co-immunoprecipitation experiments and mutational analyses. Herein, we demonstrate that RPP1 physically associates with its cognate effector ATR1 in a recognition specific manner as a prerequisite step in the induction of the hypersensitive cell death response. The LRR domain of RPP1 mediates the interaction with ATR1, while the TIR domain facilitates the induction of the hypersensitive cell death response. Additionally, we demonstrate that mutations in the TIR and NBS domains, which exhibit loss of function for the induction of the hypersensitive response, are still able to associate with the effector *in planta*. Thus, our data suggests molecular epistasis between signaling activity of the TIR domain and the recognition function of the LRR and allows us to propose a model for ATR1 recognition by RPP1.

Introduction

Plants have evolved a multi-level innate immune system to protect them against infection by a diverse range of pathogens including viruses, bacteria, fungi, oomycetes, and nematodes. Despite the great evolutionary distance among phytopathogens, the outcome of the plant-pathogen interactions is controlled by the same principles: the ability of the pathogen to suppress the plant immune system to establish infection and the ability of plants to recognize the presence of a pathogen and to induce immune responses that restrict pathogen growth.

The first line of plant defense consists of the integral plasma membrane pattern recognition receptors (PRRs) that recognize the presence of common Pathogen-Associated Molecular Patterns (PAMPs) near the cell surface. [4, 17, 18]. Upon association with PAMPs, the PRRs activate a downstream MAP-kinase signaling cascade that culminates in transcriptional activation and generation of the innate immune responses [4, 18]. This line of defense, called PAMP-triggered immunity (PTI), is commonly suppressed by a successful pathogen in order to establish infection. To interfere with PTI, pathogens from different kingdoms of life have evolved effector proteins that are delivered into and function within the host plant cells [19-22]. The second layer of plant immunity depends on the ability of the plant to recognize these pathogen-derived effectors and trigger a robust resistance response that normally culminates in a hypersensitive cell death response (HR). While PAMPs represent conserved microbial molecules, effector molecules constitute very divergent groups of proteins and their recognition requires constant structural and evolutionary adjustment of the corresponding receptors. Effector-triggered immunity (ETI) is mediated by a large group of structurally related intracellular innate immune receptors encoded by Resistance genes (*R*-genes). Products of *R*-genes either directly or indirectly recognize pathogen effectors and induce innate immunity. The major class of R-proteins is characterized by the central Nucleotide-Binding Site (NBS) domain, and carboxyl-terminal Leucine-rich repeats (LRR). This group can be further subdivided according to the amino-terminal domain into a TIR-NBS-LRR class that has a N-terminal domain with sequence similarity to *Drosophila* Toll and human Interleukin-1 receptor (TIR) and a CC-NBS-LRR class that has a structured coiled coil domain (CC). Currently, the data supporting molecular functions of the TIR/CC, NBS, and LRR domains in R-proteins is sparse, sometimes contradictory, and scattered across many different proteins and experimental systems. Therefore, there is a need for *in planta* data that would test and clearly define the roles of different domains for a given R-protein.

The NBS domain of R-proteins shares sequence similarity with the mammalian cell death inducing proteins, including human apoptotic peptidase activating factor 1 (Apaf-1), its *Caenorhabditis elegans* homologue CED-4, and a large group of intracellular Nucleotide Oligomerization Domain (NOD) receptors that function in mammalian innate immunity [23]. In animal proteins, nucleotide binding and hydrolysis acts as a molecular switch that regulates signal transduction by conformational change, often leading to oligomerization of the protein [24]. In plants, the NBS domains of the tomato R-proteins I-2 and Mi-1 were shown to bind and hydrolyze ATP *in vitro*, supporting the

function of the NBS as a nucleotide binding switch in R-protein activation [25]. In addition, mutations in the conserved P-loop motif resulted in a loss of nucleotide binding *in vitro* [25], and the corresponding loss of HR induction *in planta* [25, 26]. The intact P-loop motif was also required for effector-induced oligomerization of the tobacco TIR-NBS-LRR resistance protein N [27]. Based on this data, nucleotide binding is considered to be a molecular switch that regulates activity of R-proteins [23].

The TIR/CC N-terminal domains of R-proteins have been originally proposed to act in downstream signaling, rather than initial perception of the effector molecule. This view was based on the signaling function of TIRs in the Toll receptors from *Drosophila*, human Interleukin-1 receptor, and mammalian Toll-Like Receptors (TLRs). In the case of R-proteins, this model was supported by Frost *et al.* who showed autoactivation of the TIR domain of flax L10 and by Swiderski *et al.* who demonstrated that the TIR domains of several R-proteins, including *Arabidopsis* RPS4 and RPP1A, as well as tobacco N show different degrees of effector-independent HR when transiently expressed in *Nicotiana tabacum* or *Nicotiana benthamiana* [28]. However, there also exists evidence of TIR involvement in determining interactions with various effectors [29, 30]. Importance of the TIR domains for R-protein function was highlighted in several studies, which identified mutations in conserved residues that led to loss of HR *in planta* [28, 31].

The LRR domains are often involved in ligand-receptor interactions in animal systems [32, 33]. In plant innate immunity this role of LRR domains is supported, in the case of some R-proteins, by both genetic analyses and yeast two-hybrid data [34-36]. Evolutionary analyses suggest that many *R*-genes and their cognate effectors are evolving under the pressure of diversifying selection [15, 37]. This often results in numerous duplications and rearrangements at the *R-gene* loci as well as in high levels of polymorphisms in the LRR domain [30, 38]. Additional domain swapping experiments confirmed the role of the LRR domain in determining specificity towards different effector gene variants [30, 39-42]. Direct protein-protein interaction of the LRR domain with the cognate effector was demonstrated for rice Pi-ta and rice blast effector AvrPita. In this case, the LRR region was both necessary and sufficient for interaction in yeast two-hybrid assay and for *in vitro* binding [43]. Despite growing evidence for the direct interactions between R-proteins and their cognate effectors, there has been a lack of *in planta* biochemical data supporting the role of the LRR region in effector recognition.

The *Recognition of Peronospora parasitica* 1 (*RPP1*) locus, which contains a complex resistance gene cluster, was originally identified in *Arabidopsis thaliana* ecotype Wassilewskija (Ws) [44]. Several members of the RPP1 gene family have been shown to specify disease resistance against *Hyaloperonospora arabidopsidis* (previously known as *Peronospora parasitica*) [44-46], including RPP1-WsA, RPP1-WsB, RPP1-WsC, and RPP1-NdA, while RPP1-like genes from other ecotypes have been implicated in hybrid incompatibility [47]. Currently, only two RPP1 alleles have been shown to recognize the cognate effector *Arabidopsis thaliana recognized 1* (ATR1) from *H. arabidopsidis* [45]. One of those alleles was cloned from the *A. thaliana* ecotype Ws, and is denoted RPP1-WsB, the other, RPP1-NdA, is from *A. thaliana* Niederzenz. The R-protein RPP1-WsB and RPP1-NdA share a common TIR-NBS-LRR domain architecture, and are 87% identical at the amino acid level. Although polymorphisms are present

throughout their coding sequences, most of the differences occur in the LRR region, and include both single amino acid polymorphisms and short insertion/deletions.

ATR1 belongs to a simple locus in *H. arabidopsidis* with diverse allelic variants present in different strains of the pathogen [45]. The ATR1 protein has two regions found in most oomycete effectors thus far, including the amino-terminal eukaryotic signal peptide that targets the effector for secretion, followed by the putative host-targeting domain defined by the conserved RXLR motif [45, 48]. The rest of the protein lacks sequence similarity to any proteins of known function. Although this region is thought to contribute to the pathogen's virulence [46], recognition by RPP1-WsB and RPP1-NdA is its only known function.

In this study, we characterize the molecular mechanisms underlying the recognition of ATR1 by RPP1. We present biochemical evidence for *in planta* association between RPP1-WsB and ATR1, which correlates with the ability of RPP1-WsB to recognize this effector and elicit a resistance response. Our studies reveal that this association is mediated by the LRR domain of RPP1-WsB. Moreover, we demonstrate an epistatic relationship between effector binding ability of the LRR and HR induction by the TIR and NBS domains. Finally, we show that the TIR domain of RPP1 is sufficient for inducing cell death response.

Results

The phylogeny of ATR1 correlates with its recognition by RPP1-WsB and RPP1-NdA

Previous studies established the gene-for-gene recognition specificities of five ATR1 alleles sequenced from natural populations of *H. arabidopsidis* **(Figure 1a, Supplemental Figure 1)** and two RPP1 alleles, isolated from *A. thaliana* ecotypes Wassilewskija (Ws) and Niederzenz (Nd) **(Figure 1a, Supplemental Figure 2**) [45]. It has been demonstrated that only a subset of ATR1 alleles is able to elicit a hypersensitive response in transgenic Arabidopsis plants carrying either RPP1-WsB or RPP1-NdA. Thus, transgenic plants carrying RPP1-WsB elicited HR in response to ATR1-Emoy2, ATR1-Maks9, ATR1-Emco5, but not to ATR1-Cala2 or ATR1-Emwa1, and plants carrying RPP1-NdA recognized ATR1-Emoy2, but not ATR1-Maks9 **(Figure 1b)** [45, 46].

We decided to more closely examine the relationships between the ATR1 sequence variation and recognition by RPP1. We constructed a Maximum Likelihood phylogenetic tree of ATR1 alleles, and mapped the pattern of their recognition by RPP1. Interestingly, the three alleles recognized by RPP1-WsB clustered together, while the two non-recognized alleles formed a separate clade **(Figure 1b)**. Although the number of sequences within both ATR1 and RPP1 groups of genes is too small to make any concrete conclusions, such an evolutionary pattern is consistent with tight co-evolution between RPP1 and ATR1, and highlights the need for a deeper sequence sampling of both genes.

Allelic sequences encoding both ATR1 and RPP1 show high levels of polymorphisms. The most divergent alleles, ATR1-Emoy2 and ATR1-Emwa1, share only 81.3% sequence identity **(Supplemental Figure 1)**. Rehmany *et al.* examined the average pairwise divergence between ATR1 alleles in a sliding window analysis and found that synonymous and non-synonymous polymorphisms are distributed unevenly throughout the coding sequence, with non-synonymous mutations accumulating in the C-terminal region of the protein [45]. Similarly, sequence comparison of RPP1 alleles indicated that most of the polymorphisms accumulated in the LRR portion of the protein [44] **(Supplemental Figure 2).**

High levels of polymorphisms and evidence of positive selection in the *R-gene* and its cognate effector suggest a direct recognition of the effector protein by the corresponding resistance protein. Investigation of the molecular basis of ATR1/RPP1 interaction in a natural pathosystem presents several experimental challenges: i) low levels of protein expression for both ATR1 and RPP1 ii) small numbers of cells undergoing a resistance response during the infection and iii) inability to genetically manipulate *H. arabidopsidis*. To circumvent these problems, we adopted a transient *Agrobacterium*-mediated protein expression system in *Nicotiana tabacum*.

RPP1 activates ATR1-dependent Hypersensitive Response in *Nicotiana tabacum*

N. tabacum has been widely used for *Agrobacterium*-mediated transient expression of proteins and assaying the hypersensitive cell death response. However, one commonly occurring limitation of the heterologous plant expression system is autoactivation of the overexpressed resistance protein even in the absence of the effector. This autoactivation phenotype was observed in case of *Arabidopsis* R-proteins RPS2, RPS4 [49], and RPP13 [39], but not RPP1-WsA [50]. We tested the expression of RPP1-WsB in *N. tabacum,* driven either by its native promoter **(Supplemental Figure 3)** or the strong constitutive Cauliflower mosaic virus (CaMV) 35S promoter (35S) **(Figure 2a)**. None of the RPP1-WsB constructs exhibited autoactivation and cell death in the absence of ATR1. To find a construct amenable to biochemical analyses, we made a protein fusion between RPP1-WsB and hemagglutinin epitope tag (HA) on both N' and C' terminal ends of the protein. While protein levels of the native promoter driven RPP1-WsB construct were below detection limit, we could detect 35S:RPP1-WsB-HA by Western blot **(Figure 2b)**. The ATR1 constructs were fused with Citrine tag, a fluorescent protein variant that is detected by α-GFP antibody. Co-expression of the RPP1-WsB-HA together with either full-length ATR1-Emoy2-Citrine, ATR1Δ15-Emoy2-Citrine with truncated signal peptide sequence, or ATR1Δ51-Emoy2-Citrine truncated past the RXLR motif elicited a hypersensitive cell death response approximately 20-24 hours post infiltration **(Figure 2a)**. The full-length ATR1 protein produced a slower and less pronounced response. Western blot analysis showed that while ATR1Δ15-Emoy2-Citrine and ATR1Δ51-Emoy2-Citrine were expressed at very high level, the full-length ATR1 protein levels were below detection limit **(Figure 2c)**, probably due to partial secretion of the protein from the plant cells and instability in the extracellular

space. Based on these results, we chose to use 35S:RPP1-WsB-HA and ATR1Δ51-Citrine for further biochemical analyses.

Race-specific recognition of ATR1 by RPP1 in *Nicotiana tabacum*

In order to further characterize whether the *N. tabacum* system reflects the previously observed ATR1/RPP1 recognition specificities, we co-expressed five ATR1 alleles, ATR1-Emoy2, ATR1-Maks9, ATR1-Emco5, ATR1-Cala2, and ATR1-Emwa1 with either RPP1-WsB or RPP1-NdA. We observed the same race-specific recognition pattern in *N. tabacum* as previously described in *Arabidopsis* [45, 46], and expanded the data for RPP1-NdA, which was previously tested only with ATR1-Emoy2 and ATR1-Maks9. RPP1-WsB induced an HR when co-expressed with ATR1 alleles Emoy2, Maks9, Emco5, but not with Cala2 or Emwa1 **(Figure 3a)**; RPP1-NdA induced an HR only when co-expressed with ATR1-Emoy2, but not with any other ATR1 allele **(Figure 3b)**. The same recognition specificities were observed for both 35S:RPP1-WsB and genomic RPP1-WsB construct **(Supplemental Figure 3)**. Western blot analysis showed that while all ATR1 protein variants were expressed to equally high levels in the absence of RPP1 **(Figure 3c, left panel)**, protein levels of the recognized alleles decreased dramatically upon induction of HR **(Figure 3c, right panel)**. This data indicates that the lack of recognition of a subset of ATR1 alleles was not due to lack of protein expression. All ATR1 alleles localized consistently to the cytoplasm and nucleus, as determined by fluorescence microscopy **(Supplemental Figure 4)**, indicating that difference in recognition was not due to differential localization of ATR1 alleles within plant cell.

RPP1 Co-immunoprecipitates with ATR1

To investigate whether recognition specificity results from physical interaction between the ATR1 and RPP1 proteins, we tested ATR1 variants for their ability to associate with RPP1-WsB *in planta*. Citrine-tagged ATR1Δ51 constructs were transiently co-expressed with 35S:RPP1-WsB-HA in *N. tabacum* for 20-24 hours. Tissue was collected when the HR began to appear. Protein extracts were incubated with α-GFP antibodies to immunoprecipitate ATR1Δ51-Citrine. Resulting immunocomplexes were separated by SDS-PAGE electrophoresis, and RPP1-WsB was detected by immunoblot using α-HA antibody. Due to the mild conditions of the non-denaturing protein extraction buffer used for co-immunoprecipitations, the initial levels of RPP1-WsB protein were very low in all of the input samples and were usually close to the limit of detection by a Western blot. Therefore, we performed a parallel α-HA immunoprecipitations to ensure that RPP1-WsB protein was present in all of the samples.

We observed that RPP1-WsB co-immunoprecipitated with the three variants of ATR1 (Emoy2, Maks9, and Emco5) that also triggered HR, but did not associate with non-recognized variants ATR1-Cala2 and ATR1-Emwa1 **(Figure 4)**. This data shows a perfect correlation between the ability of RPP1 to associate with ATR1 and the activation of immune response, suggesting that it is dependent on the ability of RPP1-WsB to

associate with ATR1. Additionally, we observed a slight, but consistent decrease in the amount of RPP1-WsB associating with ATR1-Emco5 as compared to ATR1-Maks9 and ATR1-Emoy2 **(Figure 4)**.

Unfortunately, we were unable to detect robust association between RPP1-NdA and ATR1-Emoy2. In one co-immunoprecipitation experiment out of three conducted, we detected a weak signal corresponding to RPP1-NdA, which was specific for association with ATR1-Emoy2, but the level of the signal was just above the detection limit (data not shown). This could be due to a much weaker protein-protein interaction between RPP1-NdA and ATR1-Emoy2.

The LRR domain of RPP1-WsB is necessary and sufficient for association with ATR1, but not sufficient for triggering the HR

Next, we decided to determine which domains of RPP1-WsB are important for association with ATR1. We made several truncations of RPP1-WsB that carried deletions of the TIR, NBS or LRR domains **(Figure 5a).** Each truncated RPP1 construct was tagged with HA and co-expressed with Citrine-tagged ATR1Δ51-Emoy2 or ATR1Δ51-Cala2 in *N. tabacum* for 24 hours. We observed that only the full-length RPP1-WsB but none of the truncated constructs was able to elicit an ATR1-specific HR **(Figure 5b)**. We performed co-immunoprecipitation experiments to determine whether any of the RPP1 truncations were able to associate with ATR1. The results showed that the TIR or TIR-NBS constructs without the LRR domain did not co-immunoprecipitate with ATR1, while the LRR domain alone was sufficient for interactions and maintained the same binding specificity as the full length RPP1-WsB protein **(Figure 5c)**. The NBS-LRR construct showed decreased protein stability; nonetheless, we could detect a weak but specific association with ATR1-Emoy when the blot was exposed for a sufficiently long time **(Supplemental Figure 5a)**. Since the LRR domain alone co-immunoprecipitated with ATR1-Emoy2, but not with ATR1-Cala2, we concluded that the LRR domain is sufficient for interaction with the effector. Similar results were obtained from co-immunoprecipitation of the LRR of RPP1-WsB and ATR1-Maks9, ATR1-Emco5, and ATR1-Emwa1 **(Supplemental Figure 5b)**, showing that the LRR domain is sufficient for determining RPP1-WsB specificity.

Recombinant ATR1 pulls down the LRR portion of RPP1 from plant extract

In order to further define the nature of ATR1 and RPP1 interactions, we have tried expressing both proteins in exogenous systems such as E. coli and *Pichia pastoris*. While ATR1 produced good amounts of soluble protein (see chapter III), the full length RPP1 and its LRR portion not amenable to any prokaryotic or yeast protein production system that have been tested (data not included). Therefore, we examined whether the recombinant ATR1 protein produced in *E. coli* and used for crystallography, described in a subsequent chapter, was able to interact with the LRR domain produced in *N. tabacum*. We tested the ability of the recombinant ATR1 protein cross-linked to CNBr Sepharose or the CNBr beads alone to pull down LRR from plant extract **(Supplemental Figure 6a)**. We observed that recombinant ATR1-CNBr, but not

CNBr alone, was able to pull-down the LRR **(Supplemental Figure 6b)**, suggesting that ATR1 protein produced in *E. coli* maintains its functional confirmation. Additional protein bands were observed on the colloidal coomassie blue stained gel in the lanes containing ATR1, but not in the negative control (beads alone), suggesting that this pull-down system in combination with mass spectrometry could be used for the identification of additional ATR1 targets.

Mutants in TIR and NBS that do not induce HR are still able to interact with ATR1

Previous genetic analyses of R-proteins identified conserved amino acid residues in both TIR and NBS domains that were critical for the activation of HR [31, 51]. We have made a multiple sequence alignment between the TIR domains of several TIR-NBS-LRR resistance genes, including *Arabidopsis* RPP1-WsB, RPP4, RPS4, tobacco N, flax M and L, and the human Toll-like Receptor 1 (TLR1) **(Supplemental Figure 7a)**. There were several conserved residues among those R-proteins and TLR1. We used mutational analysis and identified a conserved Glutamate (E) at position 158 that was required for RPP1-WsB to elicit HR **(Figure 6b)**, while not affecting protein stability. The most conserved motif within the NBS domain is the P-loop motif (also known as kinase-1a or Walker-A), which has been shown to affect ATP binding *in vitro* and was required for eliciting *in planta* HR response [25]. The P-loop has the conserved consensus sequence GxxxxGKS/T (x is any amino acid), where the mutation in conserved Lysine (K) leads to loss of function phenotype **(Supplemental Figure 7b)** [51]. We created a K293L mutation in RPP1-WsB and found that, similar to other R-proteins, it abolished the ability of RPP1-WsB to trigger effector-mediated HR **(Figure 6b)**.

We tested the E158A and K293L mutants of RPP1-WsB for their ability to co-immunoprecipitate with ATR1. Interestingly, we discovered that both mutants retained their ability to associate with the effector **(Figure 6c)**. This data clearly shows an epistatic relationship between the recognition function of the LRR domain of RPP1-WsB and the subsequent downstream activation of signaling events mediated by the TIR and NBS domains.

The TIR domain of RPP1-WsB elicits effector-independent HR

The TIR domains of several TIR-NBS-LRR proteins, including RPP1-WsB paralogue, RPP1-WsA, are able to induce effector-independent HR in both *N. tabacum* and *Arabidopsis* [28, 52]. We were surprised that we did not observe this phenotype with our TIR-HA construct of RPP1-WsB, which was equivalent to the constructs described for RPP1-WsA. However, when we tested the same truncation variant tagged with GFP, we detected effector-independent HR which appeared around 48 hours post infiltration. Thus, the TIR domain of RPP1-WsB was sufficient for induction of the HR, yet the response was slower and weaker than effector-mediated HR induced by the full-length RPP1 protein **(Figure 7a)**. The TIR-NBS-GFP protein did not activate effector-independent hypersensitive response **(Figure 7a)**, which could be either due to decreased protein levels of this protein fusion compared to the TIR-GFP **(Figure 7d)** or

may indicate additional negative regulation of the TIR by the NBS domain. Additionally, in the case of RPP1-WsB, HR induced by the TIR truncation was dependent on the presence of a large fusion tag, the GFP. The effect of GFP could be attributed to either stabilization of the fusion protein or its native ability to form weak dimers [53]. The dimerization property of GFP can be disrupted by a single amino acid substitution A206K [53]. We tested whether dimerization of GFP might contribute to the ability of TIR to elicit autoactive HR. We observed that the TIR fusion to monomeric GFP (mGFP) was unable to elicit the response **(Figure 7b)**. This finding suggests a requirement for TIR domain dimerization prior to triggering the HR. Western blot analysis showed that the protein expression of the TIR-mGFP was slightly reduced compared to the wild type GFP version **(Figure 7e)**, therefore we cannot exclude a possibility that the ability of the TIR to induce effector-independent response depends on the relative protein stability. Additionally, we tested the E158A point mutation and observed that it abolished the ability of the TIR domain to trigger effector-independent HR **(Figure 7c)**. Western blot demonstrated that the E158A mutation did not compromise protein stability **(Figure 7f)**; therefore, loss of function in this case is not due to reduced protein expression. Collectively, this data demonstrates that the TIR domain of RPP1 is capable of eliciting HR response independently of either nucleotide binding ability of NBS or effector binding ability of LRR.

Subcellular localization of the autoactive TIR fused to GFP

To access the localization, the TIR domain (amino acids 1-266) was fused in frame with a GFP His6 tag, and expressed in *N. tabacum* under native promoter. Localization of TIR-GFP was first assessed by confocal microscopy and compared to the GFP alone. While GFP was diffused throughout both cytosol and nucleus, TIR-GFP had a unique localization pattern. The most apparent was its accumulation within the nuclei of plant cells, and absence from the cytoplasm **(Figure 8a)**. In addition, fluorescence was observed within small vesicles, accumulated around the nuclei or sometimes connected to them with fluorescent tails **(Figure 8b)**. Those vesicles did not have any apparent association with chloroplasts or Endoplasmic Reticulum. Similar pattern of TIR localization was observed in transgenic *A. thaliana* plants, containing DEX-inducible TIR GFPHis6 construct **(Figure 8b)**. Nuclear localization of TIR-GFP was also confirmed by biochemical cell fractionation. Although most of the protein was present in a soluble fraction, about 10%-20% was detected in a nuclei-enriched fraction, confirming the microscopy results **(Figure 8d)**.

The TIR-GFP-His6 fusion produces a protein with a molecular mass near the established cutoff for plant proteins to be able to passively diffuse inside the nucleus. Therefore, we tested whether TIR-GFP might be imported into the nuclei by active transport, involving nuclear import machinery. Importin α is one of the most prominent karyopherins that directly binds its cargo and facilitates nuclear import. *In planta* interactions between the TIR domain of RPP1 and importin α were tested by co-affinity purification, using a TIR-GFP-His6 fusion construct or GFP-His6 alone and antisera against native levels of importin α. *N. tabacum* extracts expressing either TIR-GFP-His6 fusion or GFP-His6 alone were incubated with with nickel–nitrilotriacetic acid agarose

(Ni-NTA). After most of the non-specific interactions were washed away, remaining protein complexes were eluted with imidazole, and analyzed by immunoblotting. All input protein extracts contained similar levels of importin α, but it interacted only with TIR-GFP-His6, but not GFP-His6 **(Figure 8c)**. This data suggested that the TIR domain of RPP1 was actively imported inside plant nuclei utilizing the classical importin α nuclear import machinery.

RPP1 contains a functional Nuclear Localization Signal Sequence

Interaction with importin α is mediated through Lys/Arg rich motifs called Nuclear Localization Sequence (NLS). The NLS in RPP1 was first predicted using the WoLF PSORT program [54]. The analysis suggested the presence of two putative bipartite NLS motifs, KRKATNQDVDSESRKRRK starting at amino acid 10 and RKTILSHILESFRRKGI starting at amino acid 97 from the amino terminus **(Figure 9a)**. To improve the accuracy of the prediction, we applied a more stringent program PredictNLS (now available through PredictProtein Server, http://www.predictprotein.org/). In addition to pattern matching-based search, PredictNLS reports the prediction accuracy, assessed by the number of experimentally confirmed nuclear proteins in which the motif is found. PredictNLS confirmed the presence of the first bipartite sequence, but did not detect the second one. We created mutations in either of the predicted sequences **(Figure 9b)** and tested their effect on localization of TIR, its ability to interact with importin α, and ability to induce HR in the context of autoactive form or full-length RPP1.

The wild type TIR domain and its NLS mutants were expressed in *N. tabacum* under the native promoter and fused with GFP-His6 on the carboxyl terminus. Localization of mutant constructs was compared to the wild type by confocal microscopy. The bipartite 1 mutant (bip1) displayed obvious change in protein distribution compared to the wild type **(Figure 9c)**. Most of the bip1 protein was now present in the cytoplasm, and not visible in the nuclei. In addition, there were no apparent fluorescent vesicles anywhere in the cell. The bipartite 2 mutant (bip2) was still detected within the nuclei, showing no apparent change in localization compared to the wild type **(Figure 9c)**. However, the bip2 mutation seemed to lower protein stability, as lower levels of the protein were detected by both microscopy and immunobloting **(Figure 9c,d)**.

Finally, both NLS mutants were accessed by their ability to form complex with importin α. Empty vector, wild type, and NLS mutant constructs were transiently expressed in *N. tabacum*, and subjected to co-affinity purifications. Although bip2 mutants were expressed at lower levels than the wild type, similar amounts of protein were purified from all extracts. Importin α was pulled down with wild type TIR and its bip2 mutant, but not with empty vector or bip1 mutant **(Figure 9d)**. Immunoblotting also showed that all extracts contained the same amount of importin α in the input. Finally, we tested whether the NLS of RPP1 would be sufficient to target an unrelated cytosolic protein (i.e. GUS) into the nucleus. We made an N'-terminal translational fusion between either GFP alone, wild type Bip1-GFP or mutated bip1-GFP with GUS

and assayed localization by confocal microscopy. The GFP-GUS fusion showed cytoplasmic localization and was completely excluded from the nuclei **(Figure 10)**. The wild type Bip1, but not bip1 mutant was able to relocalize GUS inside plant nuclei **(Figure 10)**, showing that this sequence is sufficient for nuclear localization. We can conclude that RPP1 contains a single functional NLS signal (bip1), which allows its interaction with importin α and subsequent import inside plant nuclei.

Interestingly, mutation in the confirmed NLS (bip1) did not affect HR either in the context of the autoactive TIR domain nor in the full-length RPP1 **(Supplemental Figure 8)**. At the same time, only bip2 (not an NLS) had an effect on HR, probably due to compromising protein stability or being a critical residue in TIR activity **(Supplemental Figure 8)**. Addition of the Nuclear Exclusion Sequence (NES) on the full length RPP1-WsB did not abolish the HR (data not shown). Therefore, re-localization of RPP1 away from the nucleus had no effect on downstream HR signaling.

Discussion

Pathogen effector recognition by plant resistance proteins is one of the most important initial events required for a successful immune response, yet the molecular events underlying recognition specificity remain enigmatic. Our genetic and molecular analyses of the oomycete effector protein ATR1 and its cognate resistance protein RPP1 provide a mechanistic insight into the perception of an oomycete pathogen inside the plant cell. In this manuscript we have demonstrated that the LRR domain of the RPP1-WsB protein is able to associate with the ATR1 protein *in planta*. Interestingly, this association does not require functional TIR or NBS domains that function in RPP1 activation and downstream signaling leading to activation of disease resistance responses. In this study we provide *in planta* data supporting the direct recognition of a pathogenic effector by a TIR-NBS-LRR type resistance protein.

Evolutionary histories of ATR1 and RPP1 suggest direct recognition

Recognition mechanisms of pathogenic effector proteins by corresponding R-proteins have been grouped into at least three general models. The first model proposes that recognition occurs though direct binding between the effector and cognate R-protein, manifested in the “ligand and receptor” model. The direct recognition model is supported by yeast two-hybrid interaction studies between the flax R-proteins L and M and their corresponding flax rust effectors AvrL and AvrM [34, 35]. Additional yeast two-hybrid and *in vitro* interactions have been demonstrated for tomato RRS1 and *Ralstonia solanacearum* Pop2 [36] and rice Pi-ta and rice blast AvrPita [43]. The second model suggests that recognition is the result of indirect binding, in which an effector protein interacts with a host factor bound by the R-protein. This mode of recognition has been supported by interactions between the *Pseudomonas syringae* effectors AvrPto/AvrPtoB, tomato kinase Pto and the tomato NBS-LRR protein Prf [55]. The Pto protein interacts directly with both the bacterial effectors and Pfr, providing the bridging factor in effector recognition. The third mode of effector recognition is also

indirect as recognition is achieved via detection of the enzymatic activity of the effector by the R-protein. The most well described examples of this type of recognition include *Pseudomonas syringae* effector AvrRpt2, a bacterial cysteine protease that cleaves the *Arabidopsis* host factor RIN4, and this cleavage leads to activation of the *Arabidopsis* CC-NBS-LRR protein RPS2 [56]. Similarly, *P. syringae* effector AvrPphB, a cysteine protease, cleaves the *Arabidopsis* protein PBS1 [57]. Proteolytic activity of AvrPphB is indirectly detected by the Arabidopsis CC-NBS-LRR protein RPS5 [58], which leads to activation of resistance signaling. Additional examples of indirect enzymatic activation include *Pseudomonas* effectors AvrRpm1 and AvrB, which induce phosphorylation of the *Arabidopsis* RIN4 protein leading to activation of CC-NBS-LRR protein RPM1 [59].

It has been noted that the evolutionary history of the cognate effector/*R*-gene pairs correlates with their mode of recognition. The high levels of diversifying selection in the flax rust effector alleles and corresponding flax *R*-genes has been commonly explained by an arms-race between the ligand and the receptor [35, 60, 61]. Indeed, this observation is consistent with the hypothesis that evolutionary selection pressure targets particular amino acid sites in the effector and the LRR portions of a resistance protein as a result of a direct interaction between the two proteins. On the other hand, effectors with enzymatic activity that induce modifications in the host targets are evolving under balancing or purifying selection [60, 62, 63]. Generally, it has been assumed that the *R*-genes that evolved to monitor host targets of the effectors are evolving more slowly and are less prone to duplications, since their function is to be stably associated with the invariant host proteins [18, 63].

The evolutionary history of ATR1 and RPP1 matches the direct recognition model. The ATR1 effector gene is evolving under very strong levels of diversifying selection [45]. The RPP1 gene locus shows high levels of duplications, insertions/deletions and polymorphisms, which creates alleles with altered specificities, such as RPP1-WsB and RPP1-NdA. This observation is further supported by the closer examination of the interactions between the two RPP1 alleles WsB and NdA and the ATR1 alleles Emoy2 and Maks9. Since both RPP1-WsB and RPP1-NdA can recognize ATR1-Emoy2, we can conclude that both RPP1 alleles encode functional resistance gene products capable of eliciting the defense response. Similarly, since both ATR1-Emoy2 and ATR1-Maks9 are recognized by RPP1-WsB, this interaction reveals that both effector alleles are functional. However, only ATR1-Emoy2, but not ATR1-Maks9 is recognized by RPP1-NdA. Moreover, we have demonstrated that this difference in recognition can be overturned by two independent amino acid site substitutions in ATR1-Maks9 (see next chapter). Given this evidence, we conclude that it is unlikely that RPP1 recognizes an enzymatic function of ATR1. This leaves the hypothesis of either recognition by direct binding or recognition by indirect binding, which requires additional host proteins. Using similar logic, we can conclude that if binding is indirect there should be at least two different host targets of ATR1, guarded by RPP1-WsB and RPP1-NdA. We conclude that in the light of evolutionary history of ATR1 and RPP1, recognition through direct binding remains the most parsimonious model.

Activation of RPP1 is specified by *in planta* association with ATR1

The lack of published *in planta* data demonstrating that R-proteins are able to associate with effectors inside the plant cell has led researchers to question the overall validity of the direct binding hypothesis. This lack of *in planta* data for R-protein/effector association may be explained by difficulties using natural pathosystems in biochemical assays, including very low levels of protein expression and the low strength or transient nature of R-protein/effector interactions. Here we show that the pattern of *in planta* association between ATR1 and RPP1-WsB matches that of the effector recognition suggesting that the ability of RPP1 to form a protein complex with ATR1 is required for triggering the hypersensitive cell death resistance response. At the same time, we are well aware that our *in planta* results do not prove direct protein-protein binding. Such a demonstration remains challenging due to our inability to purify soluble recombinant RPP1 protein or its LRR domain in the organisms commonly used for heterologous protein production, including *Saccharomyces cerevisiae*, *Pichia pastoris* and *Escherichia coli*. It is possible that RPP1 may require additional plant-specific factors that provide protein stability, such as HSP90, SGT1 or RAR1 or yet unidentified factors. Thus, future research will include isolating the ATR1/RPP1 protein complex from *Arabidopsis* and characterizing other interacting protein complex members by mass spectrometry to elucidate the nature of the interaction between the R-protein and its effector.

Roles of different domains of RPP1 in effector recognition and signaling responses

Although many R-proteins share a common CC/TIR-NBS-LRR domain architecture, the actual biochemical role of each domain in effector recognition and downstream signaling has been subject to debate. Based on molecular genetic analyses, the LRR domain was originally proposed to mediate effector recognition; this is well supported by yeast-two hybrid studies and *in vitro* experiments [34-36]. However, the recent *in vivo* data suggested the TIR domain of tobacco TIR-NBS-LRR protein N and CC domain of the potato CC-NBS-LRR protein Rx as the main regions involved in effector recognition [29, 64]. Despite the fact that this data contradicted previous studies of N, in which the NBS-LRR portion was shown to bind the cognate TMV effector protein p50 both *in vitro* and in yeast [65], a new hypothesis has emerged postulating that the N-terminal CC/TIR domain is the main determinant of effector/R-protein interactions [66-68]. Having established the *in planta* association between the RPP1-WsB and ATR1 proteins, we were able to address which domain of RPP1 was responsible for effector recognition. Our deletion analysis showed that neither TIR, nor TIR-NBS could associate with ATR1 *in planta*. On the other hand, the LRR domain was both necessary and sufficient for interaction with ATR1 in an allele-specific manner. Furthermore, we introduced single amino acid substitutions in the TIR domain and the P-loop of the NBS domain that abolished the ability of RPP1 to trigger effector-induced HR, but did not disrupt association with ATR1. This indicates that ATR1 can associate with RPP1 it its inactive state, which is consistent with the proposed model, in which R-

protein activation is "switched on" after perception of the effector (Takken et al, 2006; Lukasik et al., 2009). This data is also consistent with the step-wise model of R-protein activation, in which the LRR domain associates with the effector before the activation of other domains and signaling is triggered (Collier and Moffett, 2009).

Although the LRR portion of RPP1 is sufficient for association with ATR1, it is not sufficient to induce the hypersensitive cell-death immune response. Mutagenesis of the TIR and NBS domains of RPP1 revealed that they control the ability of RPP1 to activate the response after association with ATR1. Involvement of the TIR and NBS domains in signaling has been previously reported for TIR-NBS-LRR proteins L10, RPS4, RPP1A, and N [28, 49, 50, 52]. The minimal region that was sufficient to induce the HR in those proteins included the TIR domain and the first twenty amino acids of the NBS domain (which does not include the P-loop motif) [28, 52]. In this study, we report that the analogous region of RPP1-WsB (amino acids 1-266) was also able to elicit effector-independent HR when fused to the large epitope tag, GFP. Intrigued by the fact that fusion to GFP was required for the induction of the cell-death, we examined whether the dimerizing property of GFP could contribute to the signaling of the TIR domain. Indeed, when we constructed a monomeric GFP and fused it to the TIR domain, the latter lost its ability to activate effector-independent cell death. This data suggests a hypothesis whereby the activation of the TIR domain is regulated by oligomerization of RPP1 upon perception of ATR1. Effector induced oligomerization of R-proteins has been proposed previously and has been supported by oligomerization of N, which was dependent on the presence of its cognate effector and the intact P-loop motif [27]. Additionally, we have determined that the autoactive TIR domain of RPP1 localizes inside plant nuclei as well as to the intracellular vesicles of unknown origin. Nuclear localization is mediated through interaction with classical import machinery complex, importin α, and depend on the Nuclear localization sequence in the N-terminal region of RPP1. Interestingly, the exact role of the NLS in effector recognition and RPP1 activation remains elusive since mutations that abolish nuclear localization of the TIR domain of RPP1 do not affect induction of HR in *N. tabacum*.

We do not attempt to propose that our data explains how all of the R-proteins would be activated, nor that functions of the structurally similar domains would be exactly the same in different R-proteins. Different modes of effector perception predict the existence of different modes of R-protein activation. Indeed, if a protein is negatively regulated by another host factor, as in the case of *Arabidopsis* RPS2 and RIN4 [56], the initial mechanism of its activation might be different from a protein that is autoinhibited due to intramolecular interactions as demonstrated for potato R-protein Rx [69, 70]. Herein, we provide the missing biochemical *in planta* data that supports a particular simple model, in which effector recognition is mediated through the LRR domain, while the TIR domain is involved in signaling.

RPP1 activation model

Given this data, we propose the following series of molecular events that lead to activation of RPP1-WsB in response to ATR1 **(Figure 11)**. First, the pathogen infects the host, and delivers ATR1 across the haustorial extracellular matrix into the plant cell.

Recognition of ATR1 ultimately depends on which allele of ATR1 is delivered and which allele of RPP1 is present in the host. The LRR domain of RPP1 monitors for the presence of ATR1 inside the plant cell. Upon successful association between the LRR and the ATR1 variant, RPP1-WsB undergoes a conformational change, possibly oligomerization coupled to ATP binding by the NBS domain. This activates the TIR domain, which initiates a signal transduction cascade that leads to the HR and expression of plant disease resistance.

Several key questions will need to be resolved before we can claim full understanding of the RPP1/ATR1 interaction: i) the composition of the protein complex that is formed form, ii) the structural basis of effector recognition by the LRR domain, iii) the mode of TIR domain activation by LRR and NBS domains, and iv) the signaling cascade leading to the induction of the hypersensitive cell death response and plant disease resistance. The elucidation of these molecular mechanisms would be of great benefit to the field of both plant and animal immunity and will ultimately allow researchers to rationally design broad-spectrum R-genes for applications in agriculture.

Materials and Methods

Multiple Sequence Alignment

Alignment of the amino acid sequences was done using MUSCLE algorithm [71], and visualized with *belvu* [72]. For phylogenetic analyses, we used an alignment between coding DNA sequences, in which codons were aligned corresponding to the amino acid sequence alignment using *pal2nal* algorithm [73].

Phylogenetic Analysis

Phylogenetic analysis was performed using Phylip3.66 software package (Felsenstein, 2005). The unrooted maximum likelihood tree was constructed from the nucleotide sequence alignment discussed above using the *dnaml* algorithm [74] with default parameters (Supplemental Table 2). Bootstrapping was performed for 1000 replicates using the *seqboot* algorithm with the default parameters and the *consense* algorithm with the user-tree option (Supplemental Table 2). The tree was visualized using TreeView program [75].

Strains and Growth Conditions

Escherichia coli DH5α and *Agrobacterium tumefaciens* G3101 were grown in Luria–Bertani media supplemented with the appropriate antibiotics at 37 ^{o}C and 28 ^{o}C respectively. Bacterial DNA transformation was conducted using chemically competent *E. coli* (Invitrogen) and through freeze/thaw transformation of $CaCl_2$ competent *A. tumefaciens* [76]. *Nicotiana tabacum* variety Turk were grown in a controlled growth chamber at 24 ^{o}C at 16 hour light/8 hour dark photoperiod before infiltrations and switched to 24 hour light after infiltrations.

Constructs

The sequences for all of the primers used in this study are shown in Supplemental Table 1 in 5' to 3' orientation, restriction sites are indicated in bold, sequence encoding for the HA tag is underlined, pENTR/D-TOPOtargeting sequence CACC is shown in the forward primers. All point mutations were introduced by site directed mutagenesis using the Quick-Change SDM Kit (Stratagene) and primers specified in Supplemental Table 1.

The open reading frame of *RPP1-WsB* was amplified via Polymerase Chain Reaction (PCR) from *A. thaliana* Ws-0 cDNA template. Cloning of the full-length *RPP1-WsB* gene was done in two fragments, taking advantage of the unique *Nde1* restriction site within the gene. The reverse primer incorporated an in frame HA epitope tag in front of the stop codon in order to create a C-terminal fusion protein with *RPP1-WsB*. The PCR products were directly subcloned in pENTR/D-TOPOvector (Invitrogen, Carlsbad, USA). Two RPP1-WsB fragments were combined in pENTRY/TOPO with a Not1-Nde1 digest. Deletion variants of RPP1 were amplified with the following primers: *RPP1-WsB* F/TIR R; *RPP1-WsB* F/NBS R; *RPP1-WsB* NBS F/*Spe1*-HA R; *RPP1-WsB* LRR F/*Spe1*-HA R (Supplemental Table 1). A *Spe1* site was incorporated in the reverse primers, which allowed creating a C-terminal fusion with the HA or GFP epitope tags by restriction digest/ligation. For *in planta* analyses, all resulting *RPP1-WsB* constructs were introduced into pEarley Gate destination binary vector pEG201 (35S promoter, N' HA tag fusion) [77] using LR clonase (Invitrogen). Sequence information for *RPP1-NdA* was acquired from Gordon *et al.* 2002. Unfortunately, we could not isolate any functional cDNA for *RPP1-NdA*. Full-length genomic sequence, including the native promoter and native terminator was sub-cloned in three fragments, using the following primers: *RPP1-NdA* fragment 1 F/R, fragment 2 F/R, fragment 3 F/R, and joined together by restriction digest/ligation. An *Xba1* restriction site was introduced at the 3' of the *RPP1-NdA* ORF by site directed mutagenesis. The sequence encoding for a 3xHA epitope tag was PCR-amplified from the pBJ36-3xVenus-3xHA vector (provided by Jeff Long, Salk Institute), column-purified, digested with *Spe1/Xba1* and ligated into *Xba1* of *RPP1-NdA*. Resulting construct was introduced into pEG301 (no promoter) using LR clonase (Invitrogen).

All *ATR1* variants were amplified from DNA template, extracted from *H. arabidopsidis* spores of the appropriate pathovar (provided by John Mc. Dowell, Virginia Tech, Thomas Eulgem, UC Riverside, and Jonathan Jones, The Sainsbury Lab). The *ATR1* allelic variants and deletions were amplified by PCR with primers indicated in Supplemental Table 1 and subcloned in pENTR/D-TOPOvector (Invitrogen). A *Spe1* site was included in the reverse primers before the stop codon to facilitate epitope tagging. The Citrine gene was cloned on the 3' end of all ATR1 alleles using *Spe1* digestion. The *ATR1* constructs were introduced in the pEG103 (35S promoter, C' GFP fusion) and pEG202 (35S promoter, N' Flag tag fusion) binary destination vector via LR recombination.

Agrobacterium-mediated transient expression

A. tumefaciens was grown in LB broth cultures (supplemented with gentamycin 50 μg/ml, kanamycin 25 μg/ml) overnight at 28 °C with constant shaking. The cultures were spun down at tabletop centrifuge at 10,000 rpm for 2 minutes. The resulting pellet was re-suspended in 1 ml of induction medium (10mM MgCl, 10mM MES, 150mM acetosyringeone). Bacterial concentrations were measured and adjusted with induction medium to OD_{600} = 0.9. Resulting cultures were pre-induced for 2-3 hr at 28 °C. For co-infiltrations, cultures carrying individual constructs were induced separately and mixed in 1:1 ratio just before infiltration. Young tobacco leaves were inoculated with *Agrobacterium* cultures using a blunt syringe.

Transient protein expression in *N. tabacum*

To detect transient protein expression in *N. tabacum*, two leaf disks (1.5 cm diameter) were collected 24-48 hours post infiltration. The samples were frozen in liquid nitrogen, and ground with a pre-chilled plastic pestle. Protein was extracted with 150 μl of the Laemmli buffer (0.24M Tris-Cl pH 6.8, 6% SDS, 30% Glycerol, 16% β-mercaptoethanol, 0.006% Bromophenol blue, 5M Urea). Samples were boiled for 5 minutes, centrifuged at maximum speed for 10 minutes in a table-top centrifuge at room temperature; supernatants were transferred into fresh tubes before analysis by SDS-Gel electrophoresis and immunoblotting.

Co-immunoprecipitations and Metal Affinity Chromatography

All constructs were transiently expressed in *N. tabacum,* using *Agrobacterium*-mediated transformation with *A. tumefaciens* strain GV3101. Leaf tissue was collected 24 hours post infiltration when the first visible signs of HR started to appear, weighed and snap-frozen in liquid nitrogen. The weight of the tissue undergoing HR was estimated based on its surface area. We have routinely used about 1 gram of tissue for each co-immunoprecipitation experiment, which corresponds to about one youngest fully expanded leaf of three to four week old *N. tabacum*. Each tissue sample was ground with mortar and pestle to a homogeneous powder in liquid nitrogen, transferred into a pre-cooled mortar containing 2 mls of the protein extraction buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Triton X-100, 0.2% Nonidet P-40, 6mM b-mercaptoethanol, 0.3 μM Aprotinin, 10 μM Leupeptin, 1 μM Pepstatin A, and a Complete protease inhibitor cocktail (Roche)), and homogenized with a fresh pre-cooled pestle until the sample was completely thawed. Resulting samples were transferred to 1.5 ml eppendorf tubes and centrifuged for 20 minutes, 14000 rpm and 4 °C. The supernatant was split into two fresh eppendorf tubes and used as input for the α-GFP and α-HA immunoprecipitations. All the steps in the immunoprecipitations were performed at 4 °C.

In order to immunoprecipitate the target protein, either 2 μL of α-GFP (rabbit, polyclonal, Abcam) or 20 μL of α-HA (mouse, clone 16B12, Roche) was added to 1 mL of

protein extract. The antibody-lysate mixture was incubated with gentle tumbling for 1.5 hrs at 4 °C. Next, 50 μL of Protein G beads (Protein G Sepharose for Fast Flow, GE Healthcare), pre-washed 3 times in the protein extraction buffer, was added to each sample, and incubated with gentle tumbling for 4 hrs at 4 °C. Then, the beads were span down for 2 min 1000 g, and washed three times with 1 mL of the protein extraction buffer, supplemented with fresh protease inhibitors and fresh β-mercaptoethanol. The protein was eluted by boiling for 5 minutes in 50 ml of the Laemmli buffer. The samples, 5 μL per lane for detecting immunoprecipitated protein, and 25 μL per lane for detecting co-immunoprecipitating protein, were separated on 10% SDS PAGE Gels or commercial NuPAGE SDS gradient 4%-12% gels (Invitrogen), transferred to nitrocellulose membrane (Fischer), and analyzed by immunoblotting, using mouse α-GFP (clone B34, Covance) and goat α-mouse-HRP (Bio-Rad) or rat α-HA-HRP (clone 3F10, Roche). All co-immunoprecipitation experiments were performed at least three times from different leaf tissue samples, and gave robust and repeatable results.

Interactions with importin α were assayed by pulling down His6-tagged construct with Nickel Agarose and probing with antisera gainst native importin α (kindly provided by Natasha Raikhel, UC Riverside). Essentially, GFP-His6 tagged constructs were transiently expressed in *N. tabacum* and protein extraction was performed as described above. The lysates were combined with the pre-washed Ni-NTA Agarose (Quagen), and incubated with gentle shaking for 2-3 hours at 4 °C. The beads were washed 3 times with protein extraction buffer supplemented with 20 mM Imidazole and fresh protease inhibitors. The protein complexes were eluted by raising Imidazole concentration to 500 mM, and analyzed by SDS-PAGE and immunoblotting.

Protein Coupling to CNBr-Activated Sepharose and In Planta Co-Immuniprecipitations

The CNBr-ATR1 sepharose was freshly prepared for each in planta pulldown experiment and used within a week. ATR1 protein was produced in *E. coli* as described above and dialyzed into the coupling buffer (0.1M $NaHCO_3$, 0.5M NaCl, adjusted to pH 8.3 with NaOH). CNBr-activated Sepharose 4B (GE Healthcare) was pre-incubated and washed with ice-cold 1 mM HCl for four times, 15 minutes each at room temperature. At the end of the last wash, the sepharose was split in half and processed further either with addition of ATR1 or buffer alone. Usually, the ATR1 was used at concentration of 10-15 mg/mL and combined with the amount of beads corresponding to 10 mgs of ATR1 per 1 mL of sepharose. The coupling efficiency was calculated to be 80% and cross-linking efficiency was 95%. The protein was coupled overnight at 4 °C with gentle rotation on the earth-quake shaker. The remaining active groups were blocked with 0.1M Tris (pH 8.0) for 2 hours at room temperature. The excess of non-covalently bound protein was washed away with coupling buffer alternating for 3-4 times with acetate buffer (0.1M NaOAc, 0.5M NaCl, adjusted to pH 4.0 with Acetic Acid) and an at the end washing with coupling buffer again. The beads were stored in Phophate Saline Buffer at 4 °C.

The LRR domain of RPP1 was transiently expressed in *N. tabacum* and whole protein was extracted using the pull-down buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Triton X-100, 0.2% Nonidet P-40, 6mM β-mercaptoethanol, 0.3 μM Aprotinin, 10 μM Leupeptin, 1 μM Pepstatin A, and a Complete protease inhibitor cocktail (Roche)) as described before [16]. The 1.5 mL of plant lysate was combined with 50 μL of either CNBr-ATR1 or CNBr alone pre-washed three times in pull-down buffer. The samples were incubated for 2-3 hours at 4 ^{o}C, washed three times with 1 mL of pull-down buffer. Bound proteins were eluted from the bead by boiling in 50 μL Laemmli buffer and analyzed by SDS-PAGE and immunoblotting.

Sub-cellular fractionations

Separation of nuclei from other cellular compartments were done according to Wirthmueller et al. 2007. Note, that this fractionation method produces crude fractions, and nuclear fraction might contain chloroplasts.

Accession numbers

Sequence data that were used in this paper can be found in the NCBI databases under the following accession numbers: *ATR1-Emoy2* (gi61660946), *ATR1-Maks9* (gi61660952), *ATR1-Emco5* (gi61660954), *ATR1-Cala2* (gi61660958), *ATR1-Emwa1* (gi61660960), *RPP1-WsB* (gi3860164). The genomic sequence of RPP1-NdA was cloned based on Gordon *et al.* 2002, and deposited in the GenBank database under the accession number HM209027 (protein id ADI80539).

Figure 1.

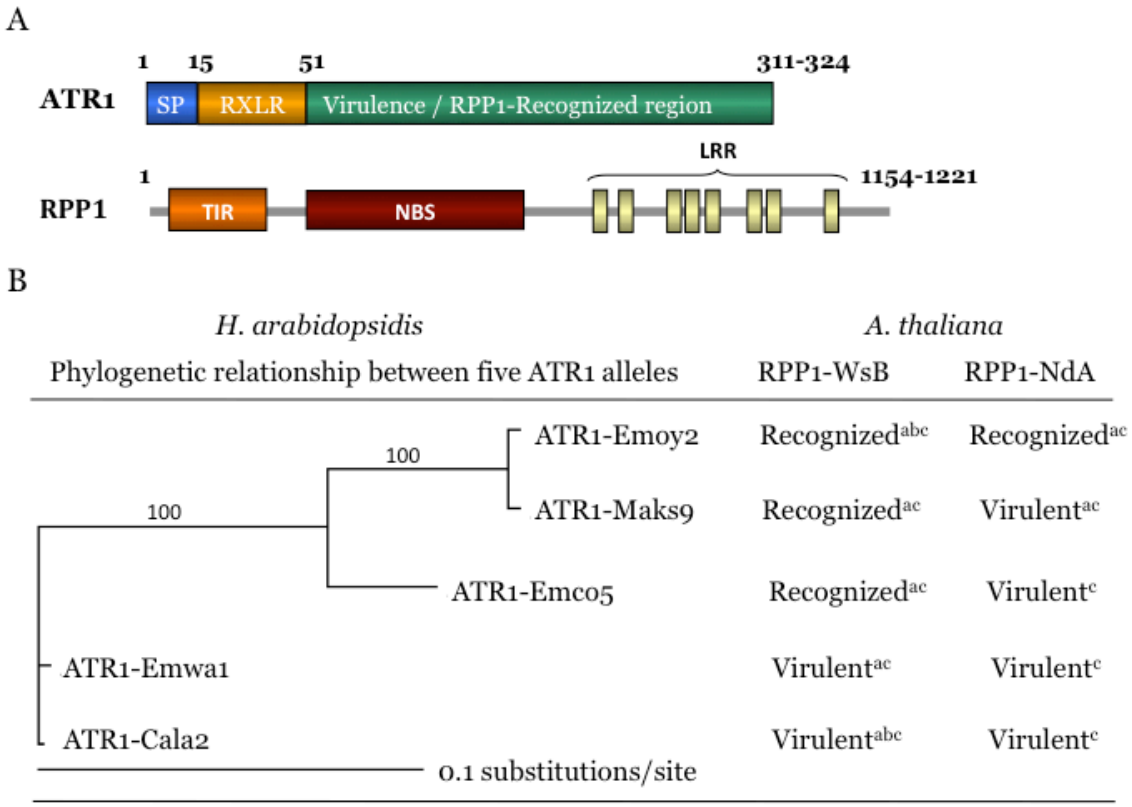


Figure 1. Domain architecture and allele-specific recognition pattern of ATR1 and RPP1.
A) Domain organization of an oomycete effector ATR1 and Arabidopsis resistance protein RPP1. The ATR1 diagram shows the eukaryotic secretion sequence (SP), host-translocation targeting region, and C-terminal domain, recognized by RPP1. The RPP1 diagram highlights the major domains identified by sequence similarity to known protein families: Toll/Interleulin 1 homology domain (TIR), nucleotide-binding site region (NBS), and series of Leucine-rich repeats (LRR). B) The maximum-likelihood tree illustrating phylogenetic relationship between five ATR1 alleles (left) and corresponding recognition pattern by RPP1 alleles RPP1-WsB and RPP1-NdA (right). Bootstrap values are shown on the tree branches. Recognition was demonstrated in the following studies: [a] in *Arabidopsis* in co-bombardment assay [45], [b] in *Arabidopsis* with TTSS delivery by *Pst* DC3000 [46], [c] in *N. tabacum* by *Agrobacterium*-mediated transient co-expression (this study).

Figure 2.

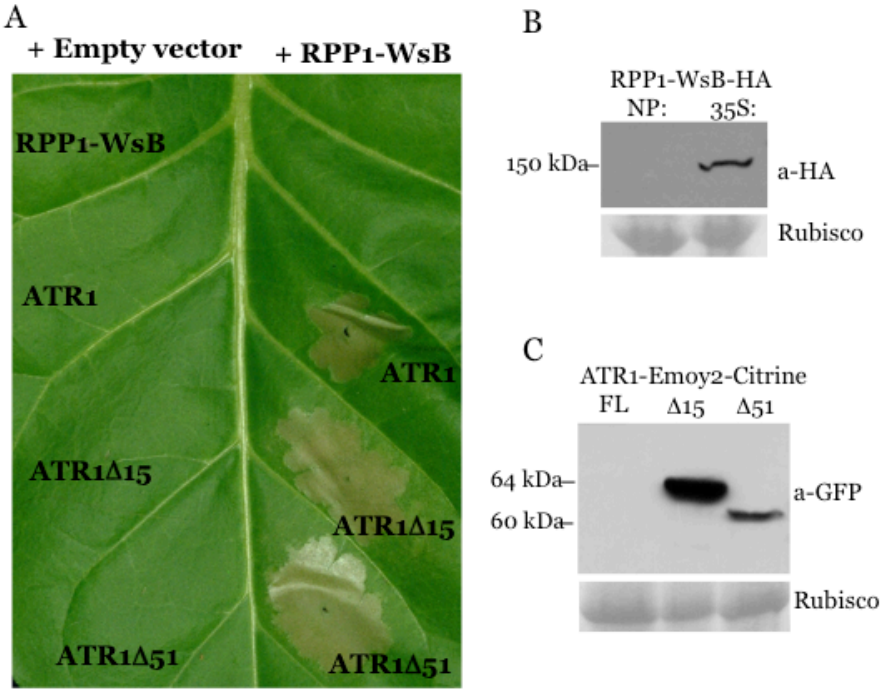


Figure 2. Co-expression of ATR1 and RPP1 in *N. tabacum* triggers the hypersensitive response
A) RPP1-WsB-HA induces HR in *N. tabacum* upon *Agrobacterium*-mediated co-expression with full-length ATR1-Emoy2-Citrine, ATR1Δ15-Emoy2-Citrine or ATRΔ51-Emoy2-Citrine. All constructs were expressed under the control of constitutive Cauliflower mosaic virus 35S promoter. The picture was taken 48 hours post-infiltration. B) Western blot (WB) showing relative levels of protein expression of RPP1-WsB-HA in *N. tabacum* driven by either its native promoter (NP) or Cauliflower mosaic virus 35S promoter (35S). Rubisco, stained on the blots with Ponceau S stain, is shown as a loading control. C) WB showing relative levels of protein expression of the full-length ATR1-Emoy2-Citrine, ATR1Δ15-Emoy2-Citrine and ATRΔ51-Emoy2-Citrine in *N. tabacum* at 24 hours post-infiltration. Rubisco – loading control.

Figure 3.

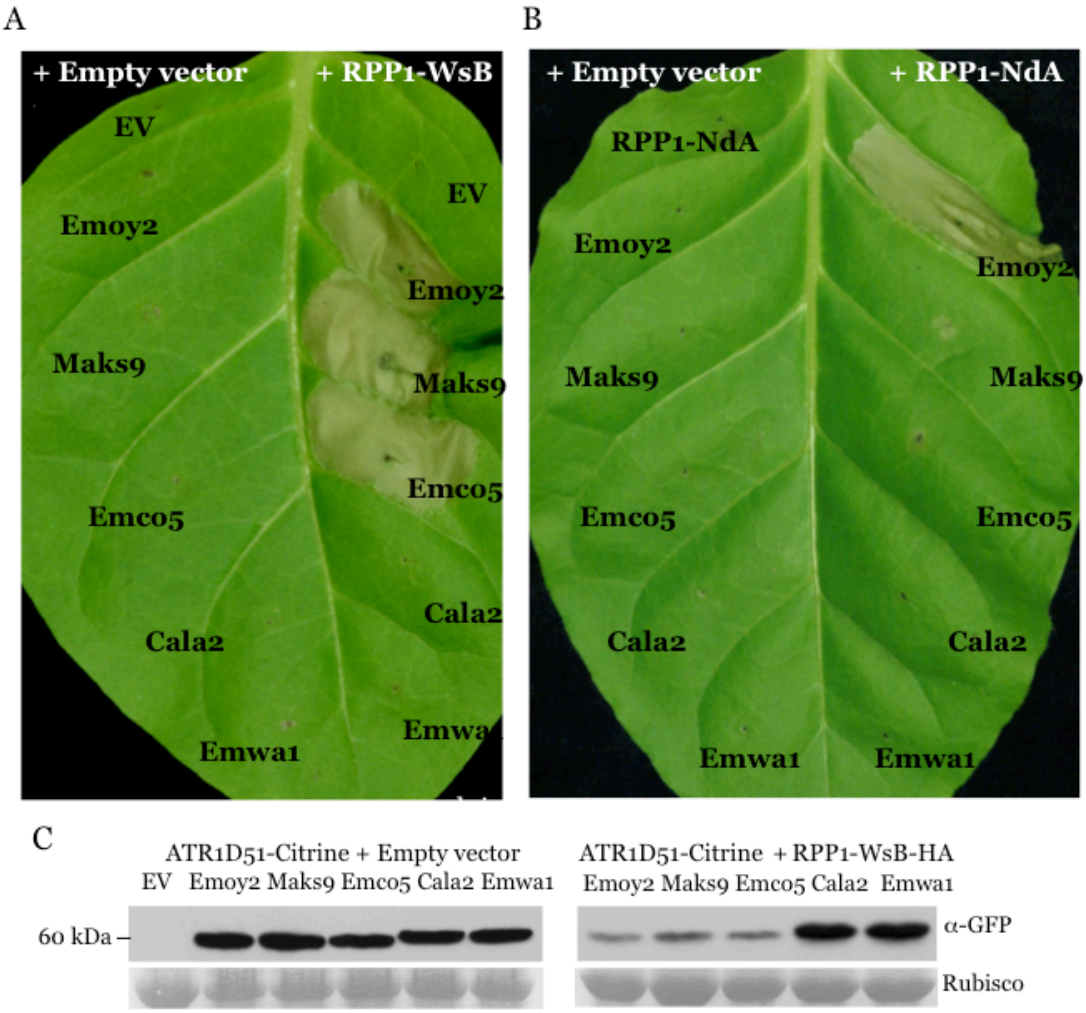


Figure 3. Race-specific recognition of ATR1 alleles by RPP1-WsB and RPP1-NdA in *N. tabacum*

(A) Co-infiltration of RPP1-WsB with ATR1-Emoy2, ATR1-Maks9, and ATR1-Emco5, but not ATR1-Cala2 or ATR1-Emwa1 triggers the HR in *N. tabacum*. *(B)* Co-infiltration of RPP1-NdA with ATR1-Emoy2, but none of the other four alleles triggers the HR. Pictures were taken 48 hours post infiltration: EV - Empty vector. *(C)* WB showing relative levels of protein abundance of five ATR1 variants co-infiltrated with empty vector (left panel) or with RPP1-WsB (right panel). Samples for protein extraction were collected at 20-24 hours post infiltration, when the HR started to appear. Rubisco – loading control.

Figure 4.

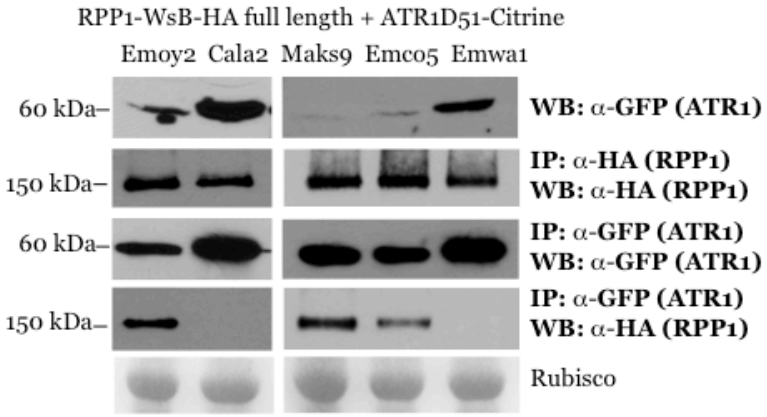


Figure 4. RPP1-WsB associates with ATR1 alleles *in planta* in a recognition-specific manner.
RPP1-WsB-HA co-immunoprecipitates with the recognized ATR1 alleles ATR1-Emoy2-Citrine, ATR1-Maks9-Citrine, ATR1-Emco5-Citrine, but not with the virulent alleles ATR1-Cala2-Citrine and ATR1-Emwa1-Citrine. RPP1 and ATR1 alleles were transiently expressed in *N. tabacum* leaves by *Agrobacterium* infiltration. Tissue samples were collected at 24 hours pos-infiltration and analyzed by immunoprecipitation and Western blot. Top panel: input probed with α-GFP for ATR1-Citrine; second panel: α-HA IP showing the presence of RPP1-WsB protein in all samples; third panel: α-GFP IP samples probed with α-GFP for immunoprecipitating ATR1; forth panel: α-GFP IP probed with α-HA for co-immunoprecipitating RPP1. Rubisco – loading control.

Figure 5.

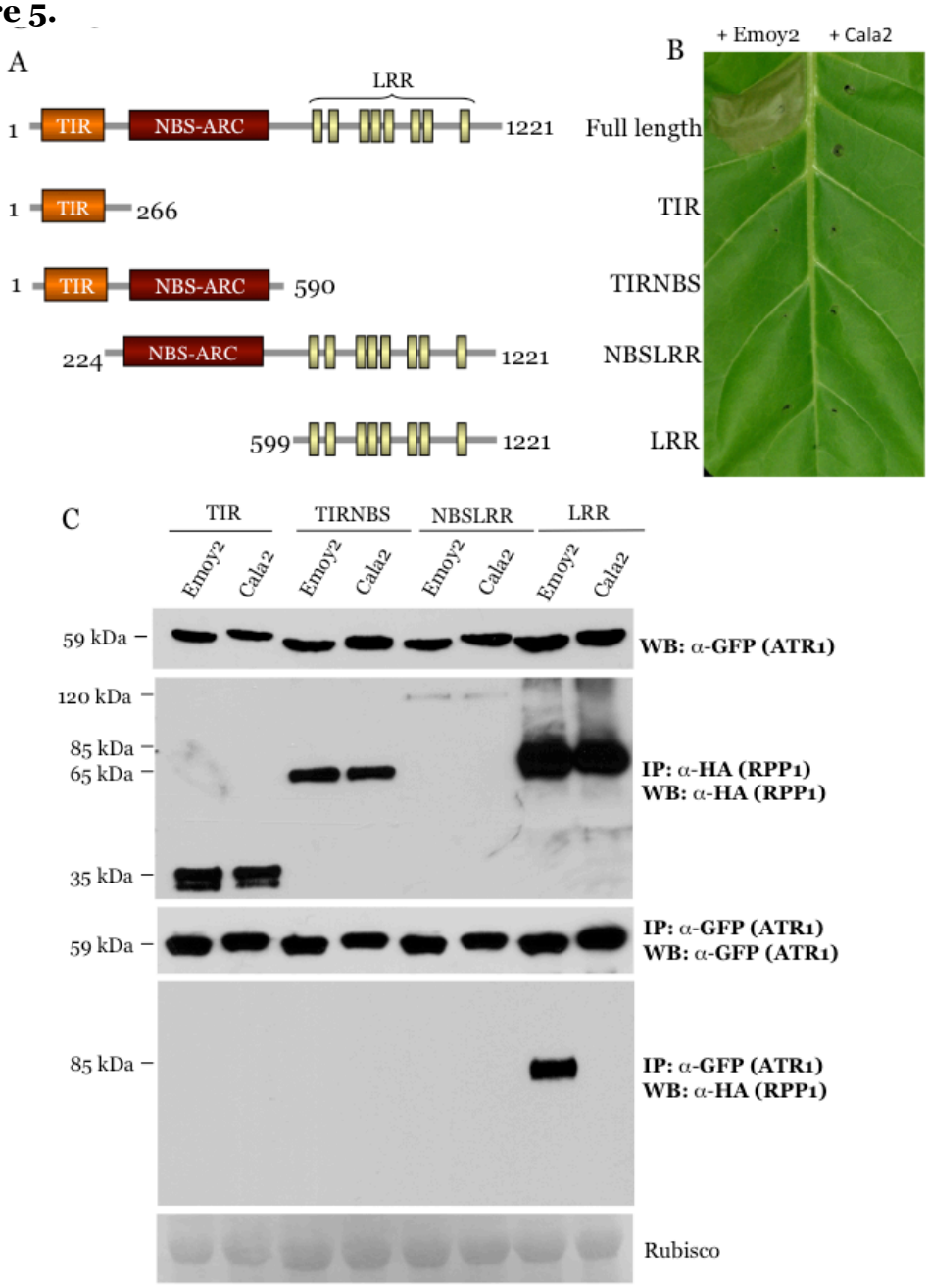


Figure 5. The LRR region of RPP1 is necessary and sufficient for association with ATR1, but not for eliciting the HR.
(A) Overview of the RPP1-WsB truncations. *(B)* Corresponding phenotypes when co-expressed with ATR1Δ51-Emoy2 or ATR1Δ51-Cala2 in *N. tabacum*. The picture was taken 48 hours post-infiltration. *(C)* Co-immunoprecipitation of RPP1-WsB truncations with ATR1. LRR domain alone, but not TIR or TIR-NBS domains associate with ATR1 *in planta*. NBS-LRR construct is unstable *in planta,* but shows allele-specific binding to ATR1 upon prolonged exposure (see Results section). Panels are labeled as in Figure 4.

Figure 6.

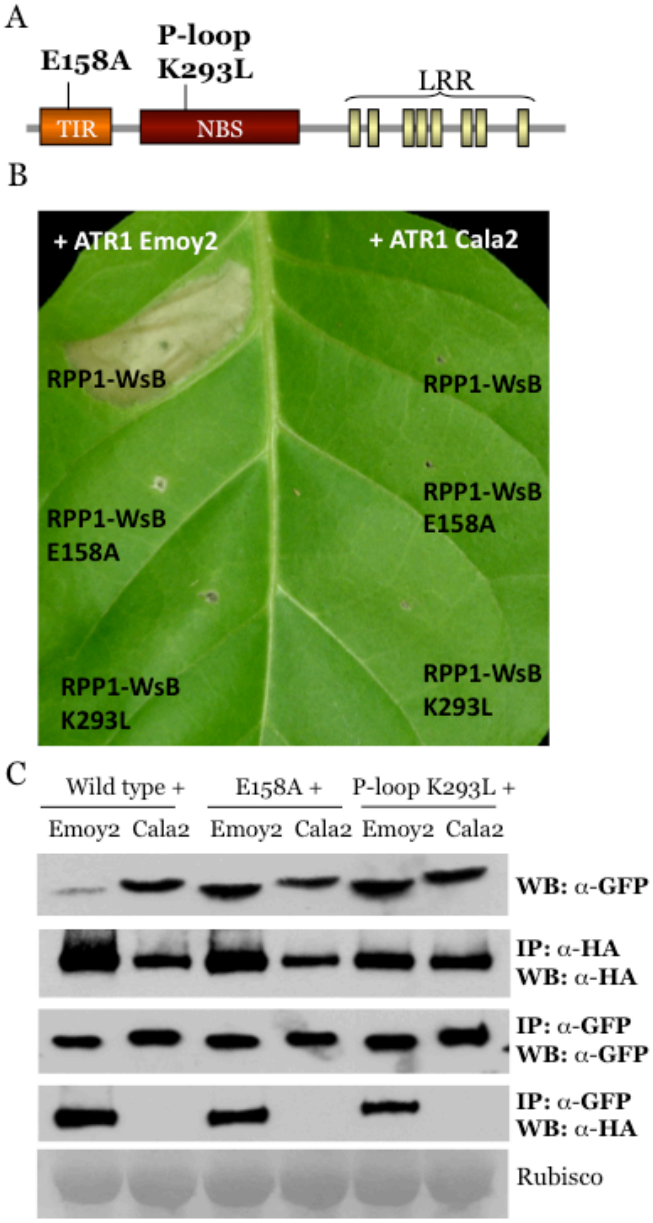


Figure 6. Mutations in the TIR and NBS domains that compromise induction of HR do not affect *in planta* association between ATR1 and RPP1.

(A) Schematic diagram of RPP1-WsB, showing the position of the E158A mutation in the TIR domain, and of the P-loop K293L mutation in the NBS domain. *(B)* RPP1-WsB-HA E158A and RPP1-WsB-HA K293L are unable to induce HR in *N. tabacum* upon co-expression with ATR1Δ51-Emoy2-Citrine. Pictures taken at 48 hours post-infiltration. *(C)* RPP1-WsB E158A and RPP1-WsB K293L are still able to associate with ATR1-Emoy2, but not with ATR1-Cala2. Panels are labeled as in Figure 4.

Figure 7.

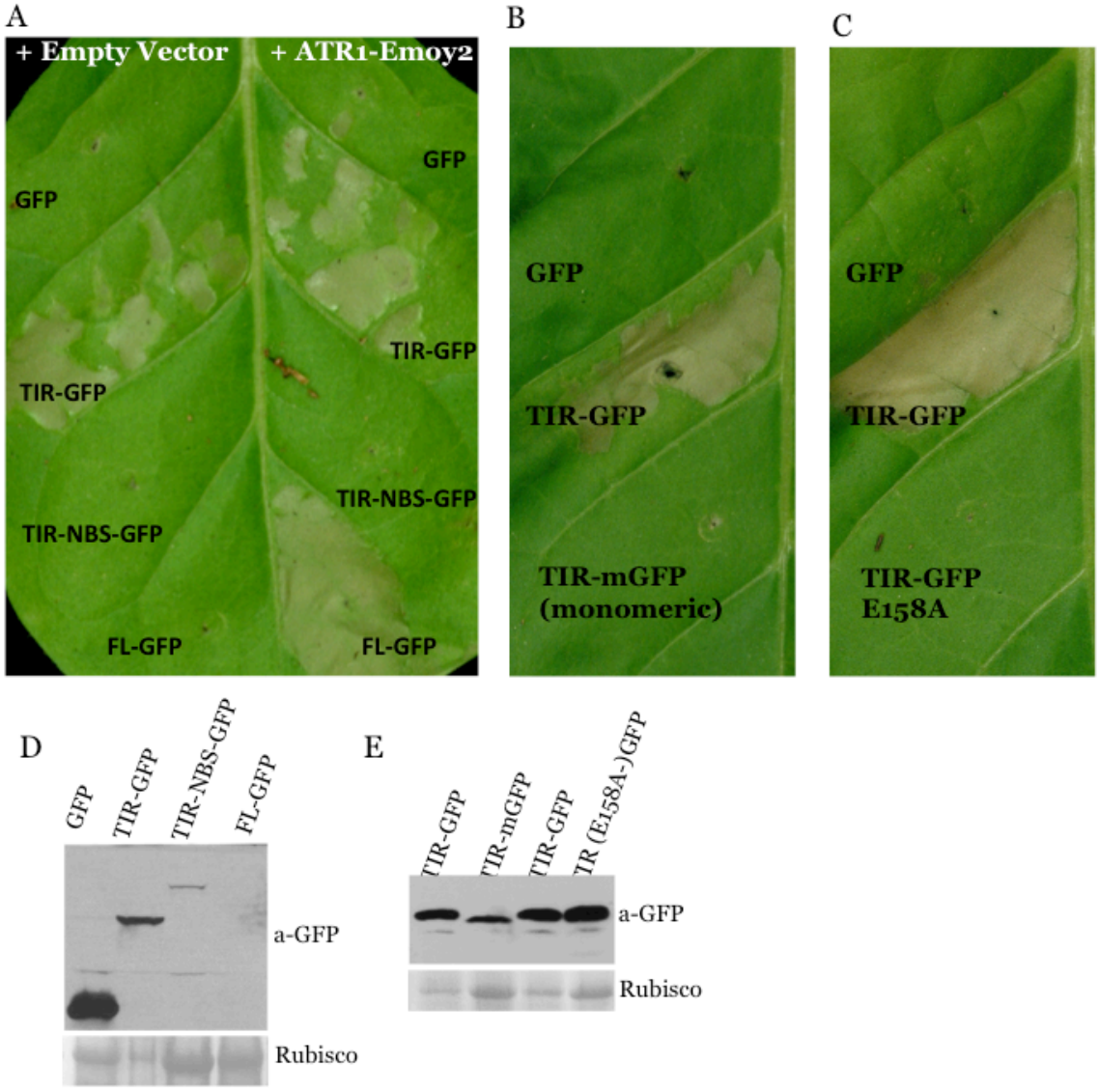


Figure 7. The TIR domain is sufficient for triggering the HR
(A) The TIR truncation (aa 1-266) of RPP1-WsB fused to GFP is able to elicit effector-independent HR in *N. tabacum*. *(B)* An A206K mutation in GFP that disrupts its ability to form dimers compromises the ability of the TIR to autoactivate. *(C)* The E158A mutation in the TIR domain compromises its ability to trigger effector-independent HR. *(D-F)* Western Blots showing relative expression levels of the GFP-tagged RPP1 constructs. Samples were taken at 48 hpi. Rubisco – loading control. At 48 hpi HR triggers overall protein degradation (ie lower levels of Rubisco), while levels of TIR-GFP are unchanged.

Figure 8.

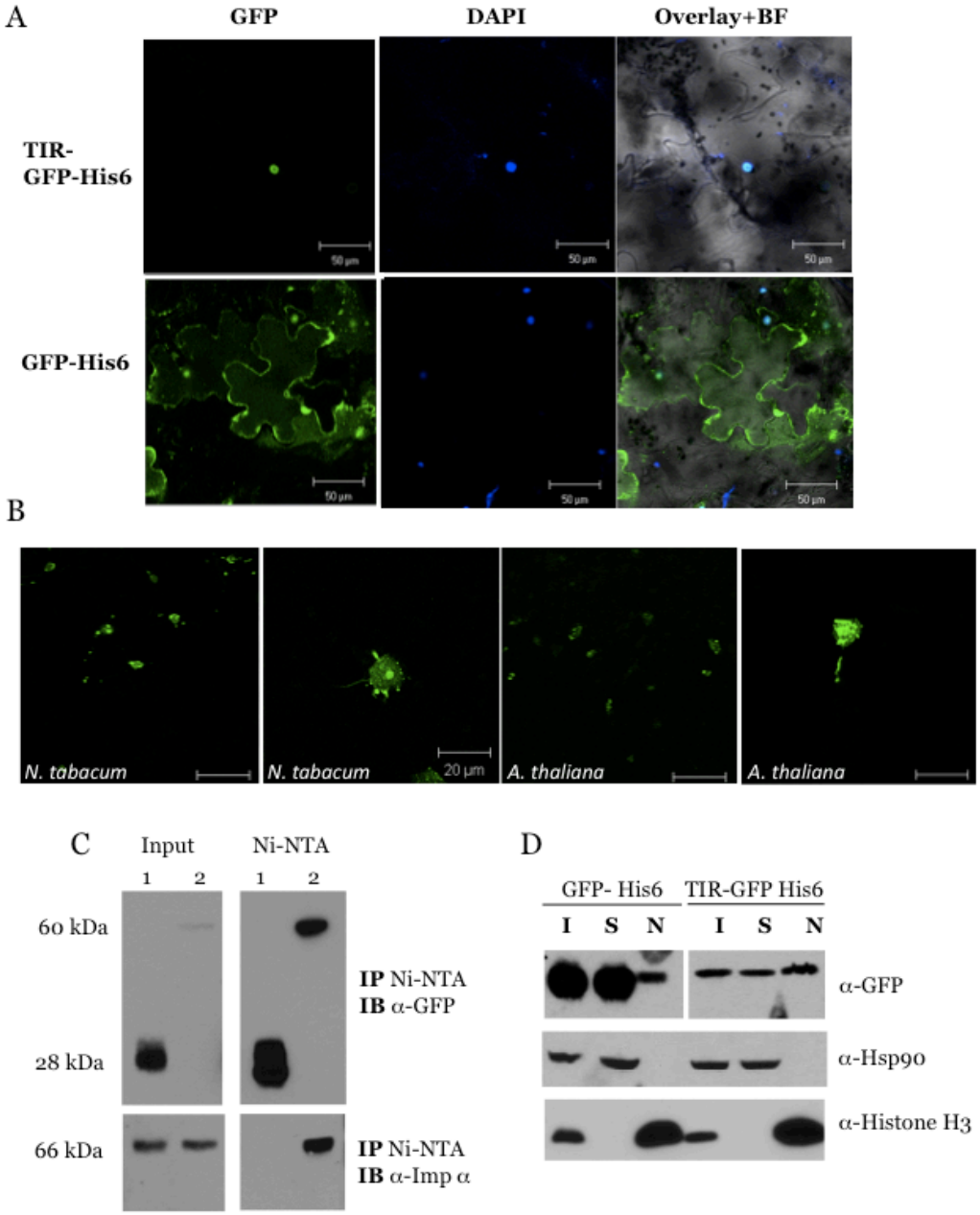


Figure 8. The autoactive TIR domain localizes to intracellular vesicles and inside plant nuclei.

(A) Confocal microscopy images showing localization of TIR-GFP-His6 (top panel) compared to GPF-His6 alone (bottom). DAPI, 4',6-diamidino-2-phenylindole, provides fluorescent staining of nuclei. BF – bright field. *(B)* Localization of TIR-GFP-His6, highlighting the intracellular vesicles, shown in *N. tabacum* and *A. thaliana*. *(C)* Pull-down of GFP-His6 (lane 1) and TIR-GFP-His6 (lane 2) to assay for interaction with importin α. *(D)* Crude sub-cellular fractionations. I – input, S – soluble, non-nuclear fraction, N – nuclei enriched fraction.

Figure 9.

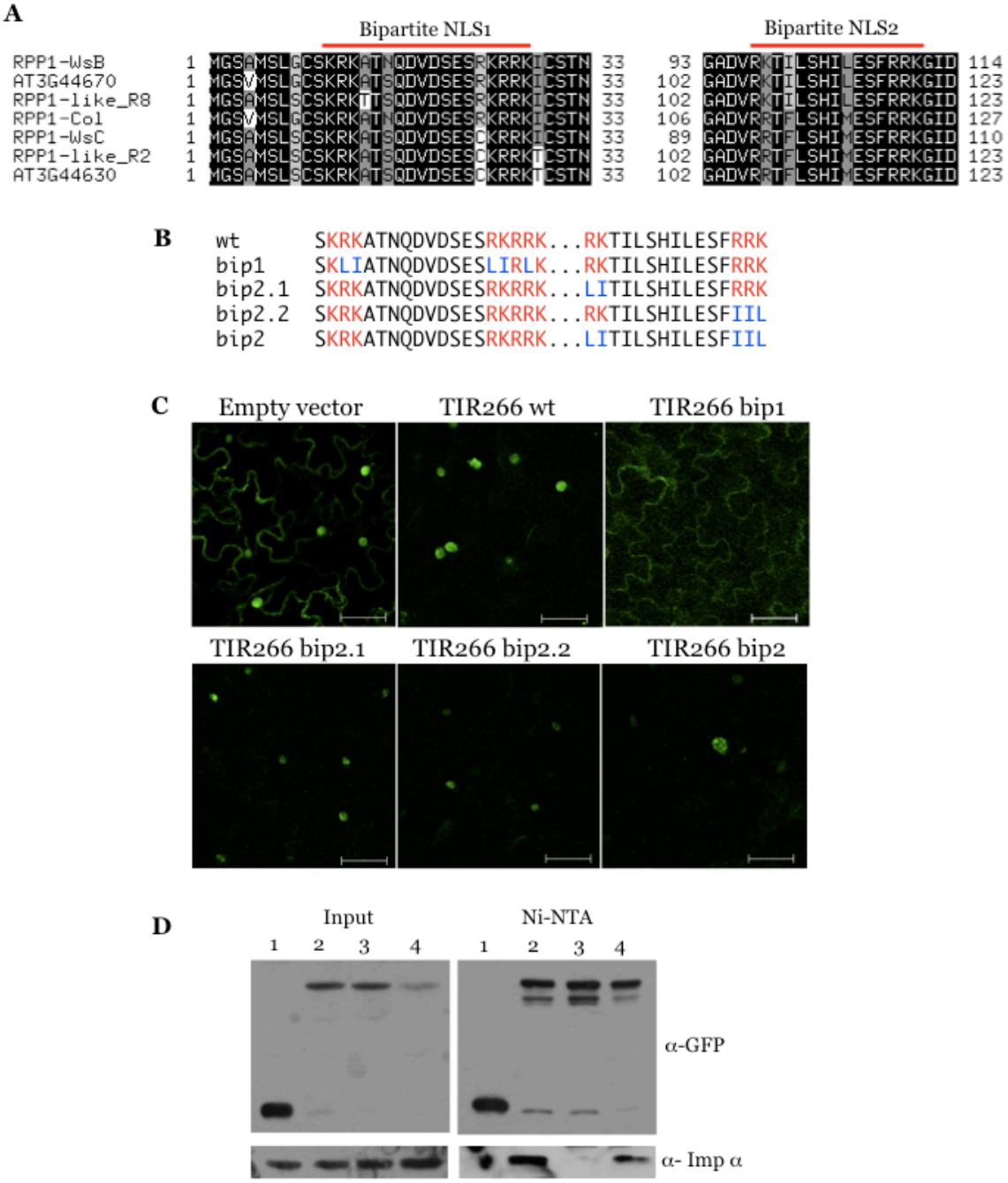


Figure 9. RPP1 contains a functional nuclear localization sequence (NLS)
(A) Multiple sequence alignment of RPP1 and its homologs containing two putative bipartite NLS, identified by bioinformatic analyses. *(B)* Set of NLS mutations (bip) used in this study. *(C)* Confocal microscopy showing localization of bip mutants. *(D)* In planta pull-down assaying for interaction with importin α. 1 – GFP-His6 alone, 2 –wild type TIR, 3 – bip1 TIR, 4 – bip 2 TIR.

Figure 10.

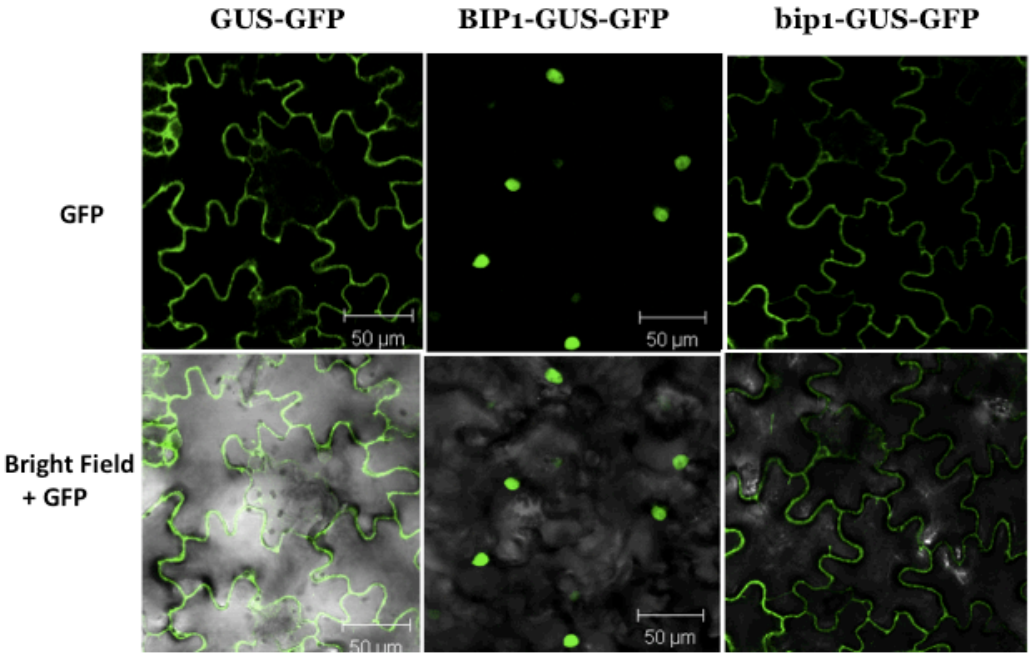


Figure 10. The NLS of RPP1 is sufficient for targeting an unrelated protein to the nucleus
Confocal microscopy showing the localization of β-glucuronidase, GUS, protein fused to GFP alone, together with the bipartite NLS sequence of RPP1 (BIP1) or its mutated form (bip1).

Figure 11.

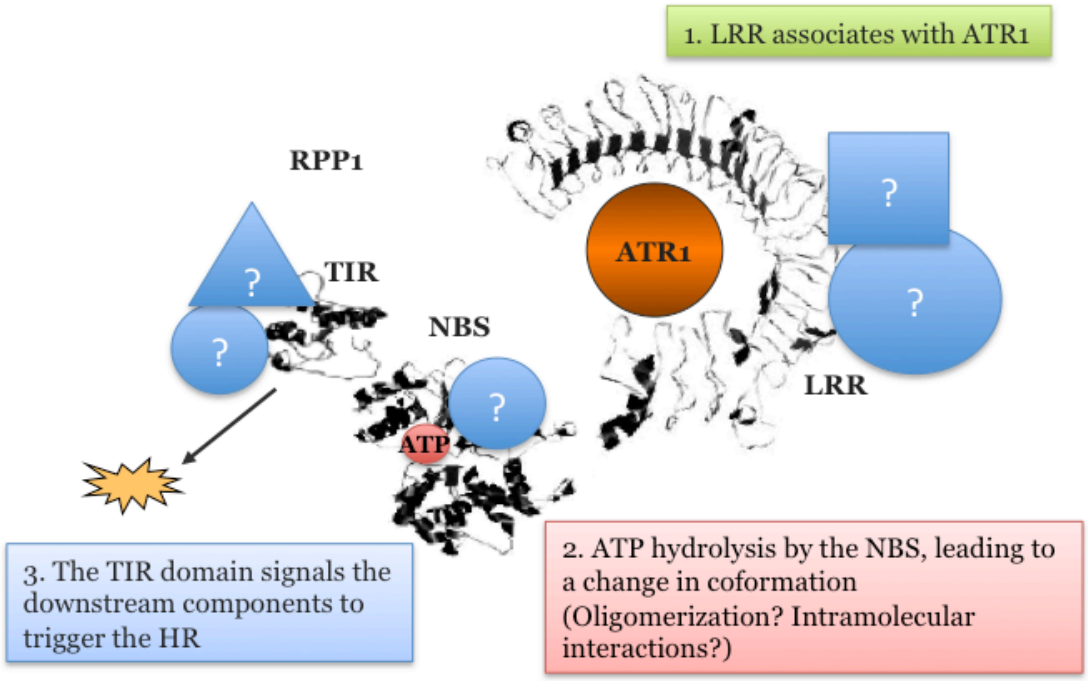
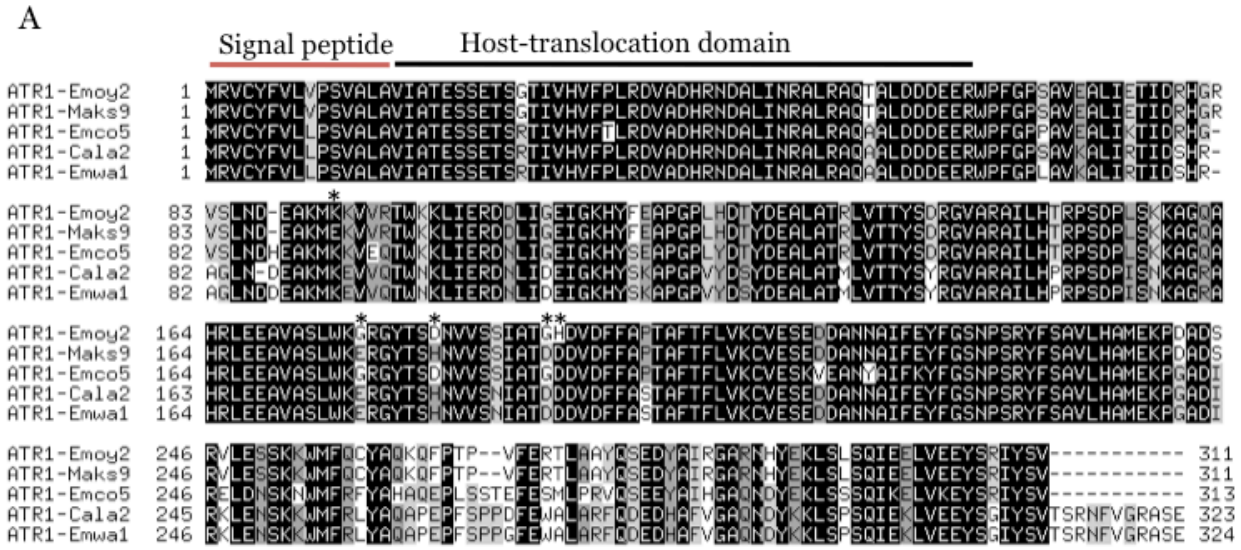


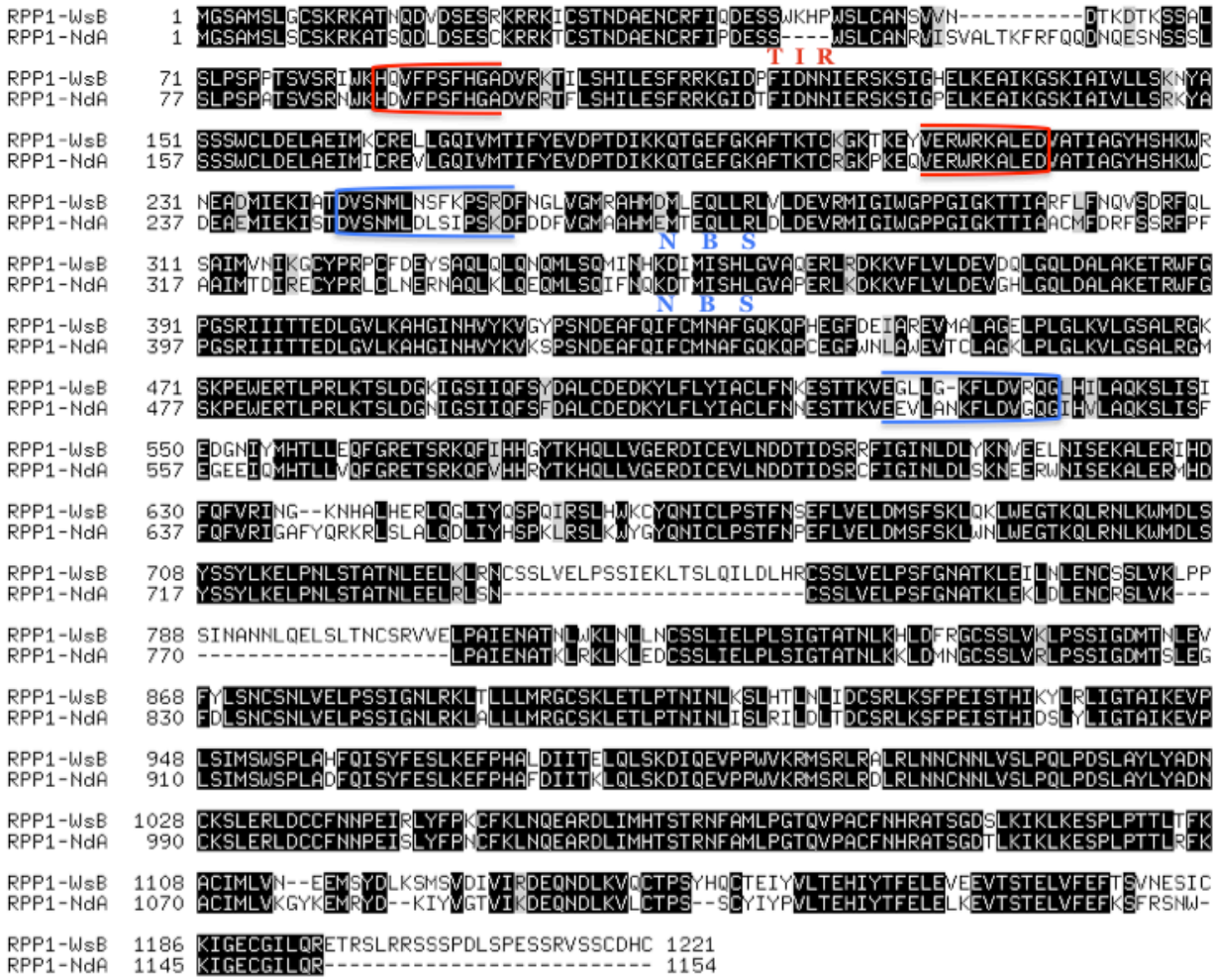
Figure 11. Model for ATR1 effector recognition and RPP1 activation
Predicted sequence of molecular events that specify ATR1 recognition and subsequent activation of RPP1. All RPP1 domains are presented as structural models. The TIR and the LRR domains were modeled after human Toll/Interleukin receptor 3, and NBS domain – after Apaf1.

Supplemental Figure 1.



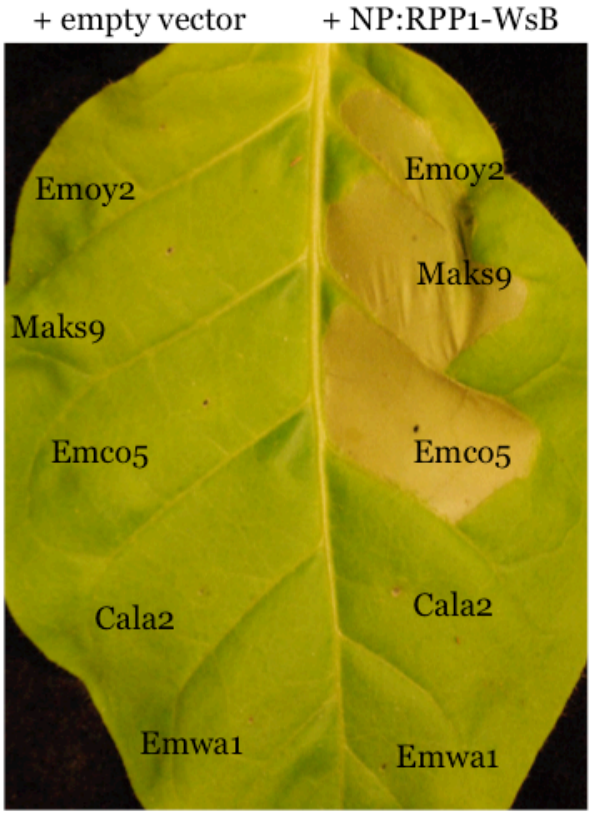
Supplemental Figure 1. Multiple Sequence Alignment of the ATR1 protein sequences used in this study.
Signal peptide and host-translocation domains are highlighted on top of the alignment. Five amino acid sites that differ between ATR1-Emoy2 and ATR1-Maks9 are marked with asterisks (*).

Supplemental Figure 2.



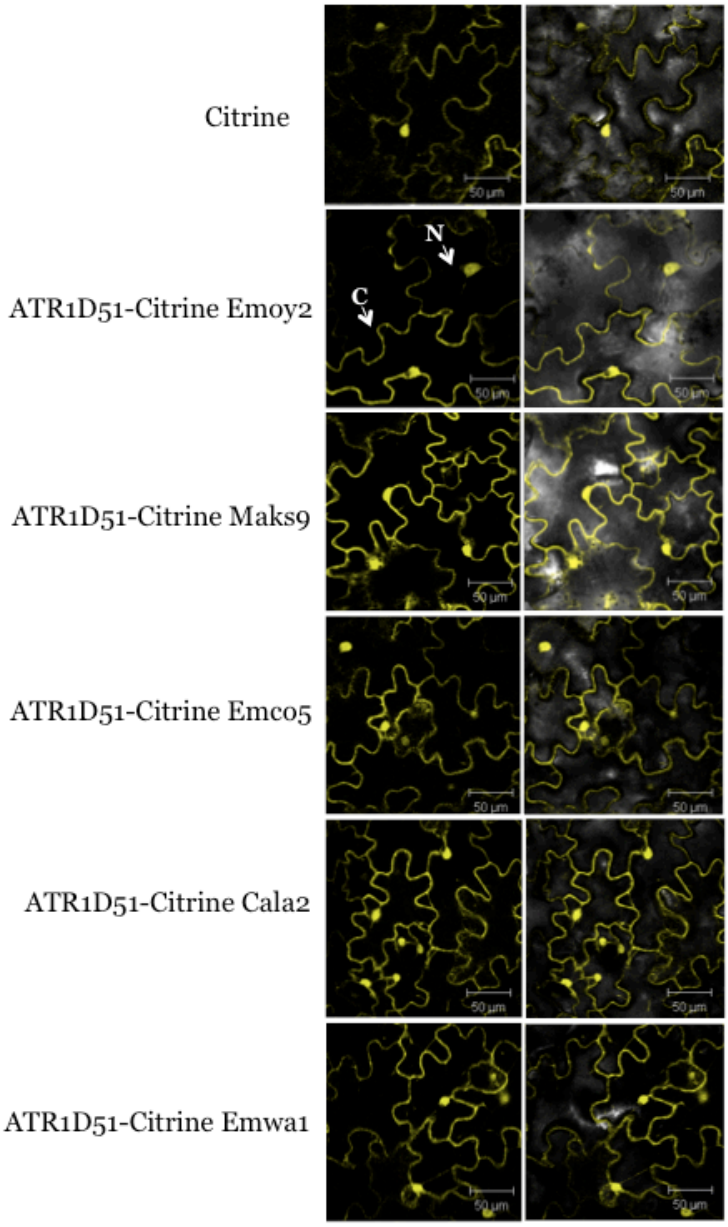
Supplemental Figure 2. Multiple Sequence Alignment of the protein sequences coding for RPP1-WsB and RPP1-NdA.
The boundaries of the TIR and NBS domains, as predicted by Pfam, are highlighted in the alignment.

Supplemental Figure 3.



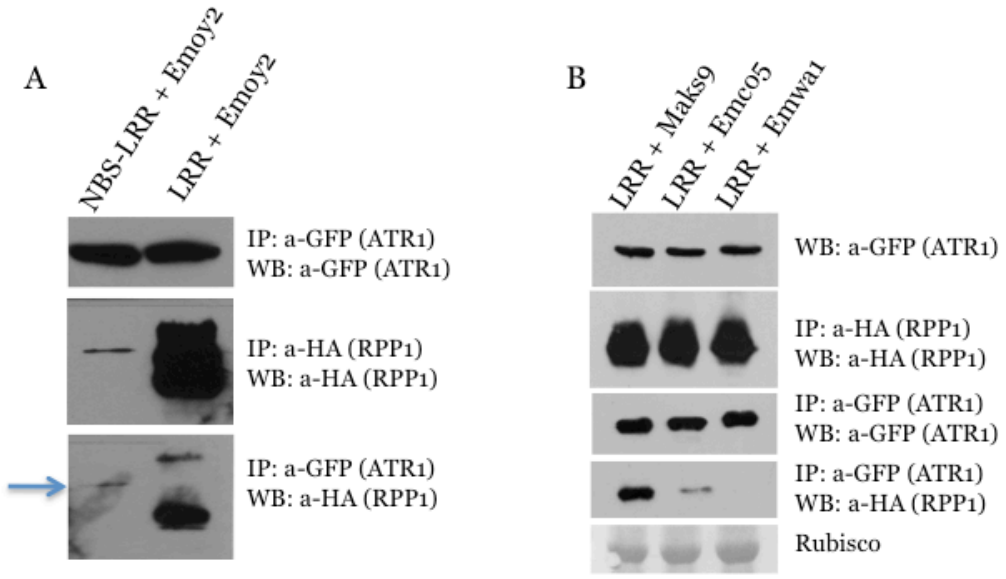
Supplemental Figure 3. Race-specific recognition of ATR1 alleles by the genomic RPP1-WsB construct expressed under its native promoter.
Co-infiltration of gRPP1-WsB with ATR1Δ51-Emoy2, ATR1Δ51-Maks9, and ATR1Δ51-Emco5, but not ATR1Δ51-Cala2 or ATR1Δ51-Emwa1 triggers the HR in *N. tabacum*.

Supplemental Figure 4.



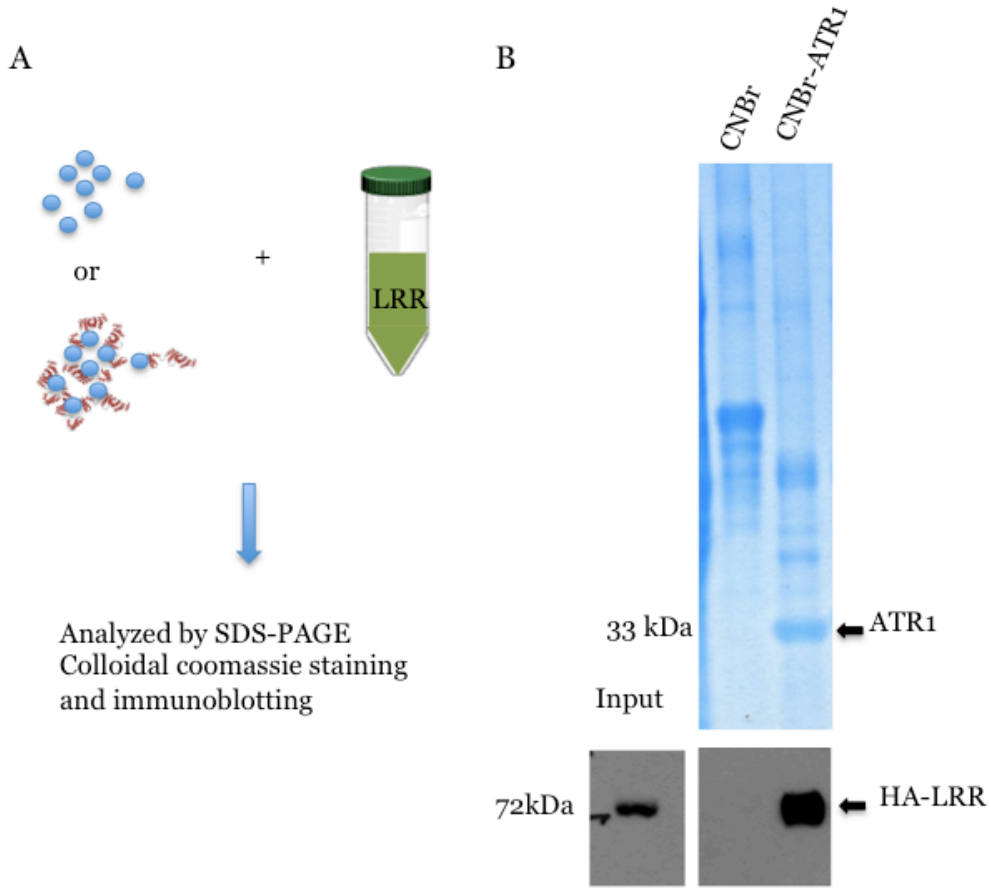
Supplemental Figure 4. Localization of different ATR1 alleles in *N. tabacum*.
Subcellular localization of ATR1Δ51-Citrine or Citrine alone was observed in live *N. tabacum* tissue, mounted in water, using confocal microscope. Fluorescence of Citrine (left panels) or bright field overlaid with Citrine (right panel) is shown. A typical view of nucleus and cytoplasm are indicated on the ATR1Δ51-Citrine Emoy2 panel. N – nucleus. C – cytoplasm.

Supplemental Figure 5.



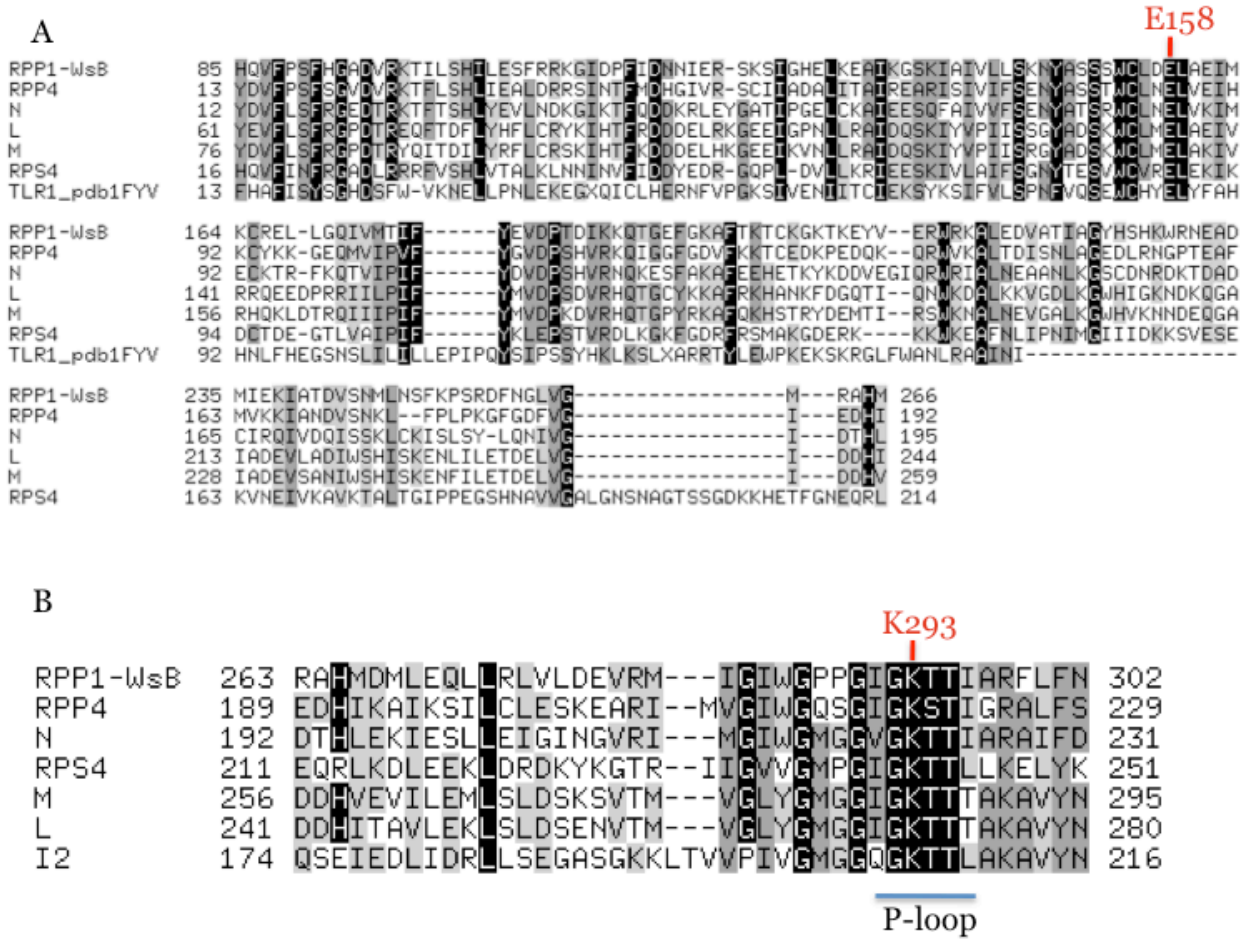
Supplemental Figure 5. Additional co-immunoprecipitation experiments showing interactions between NBS-LRR and LRR domains of RPP1-WsB and different ATR1 variants.
A) The NBS-LRR-HA of RPP1-WsB co-immunoprecipitates with ATR1Δ51 Emoy2-Citrine. B) The LRR-HA domain of RPP1-WsB immunoprecipitates with ATR1Δ51-Maks9-Citrine, ATR1Δ51-Emco5-Citrine, but not with ATR1Δ51-Emwa1-Citrine. Panels are labeled as in Figure 4.

Supplemental Figure 6.



Supplemental Figure 6. CNBr pulldowns
(A) Overview of the experiment. Beads represent CNBr-activated sepharose with ATR1Δ51-Emoy2 protein shown in red. The LRR is transiently expressed in *N. tabacum* and total protein lysate is combined with CNBr or CNBr-ATR1. *(B)* Elutions from the pulldowns separated on SDS-PAGE gel and either stained with colloidal coomassie (top) or immunoprobed for the LRR with α-HA (bottom). A 33 kDa protein band corresponding to ATR1 and confirmed by mass spectrometry is marked with an arrow.

Supplemental Figure 7.

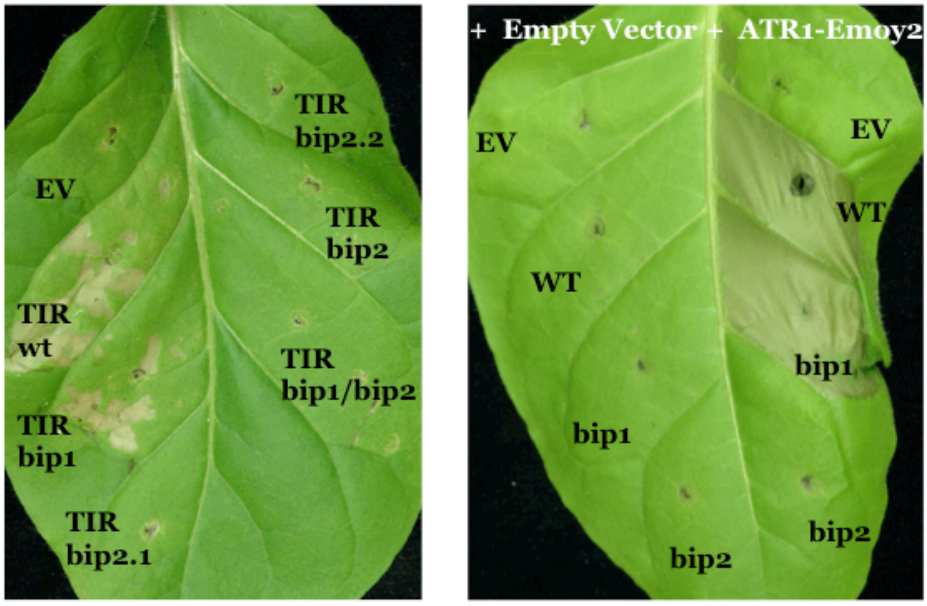


Supplemental Figure 7. Multiple Sequence Alignment (MSA) of the TIR domain and P-loop motif in different R-proteins and the TIR domain of human Toll Like Receptor 1.

A) MSA of the protein sequences of the TIR domains in *Arabidopsis* proteins RPP1-WsB, RPP4, and RPS4, tobacco N, flax M and L6, and human Toll Like Receptor 1. Alignment is colored based on the conservation levels. The conserved residue E158 (amino acid position in RPP1-WsB) used for mutational analysis in this study is marked on top of the alignment.

B) MSA of the P-loop region in NBS domains of *Arabidopsis* R-proteins RPP1-WsB, RPP4, RPS4, tobacco N, flax M and L, and tomato I-2. The P-loop motif is highlighted on the bottom of the alignment. Conserved lysine 293 (amino acid position in RPP1-WsB) is indicated on top of the alignment.

Supplemental Figure 8.



Supplemental Figure 8. Nuclear localization of RPP1 is dispensable for HR induction in *N. tabacum*

(A) Autoactive induction of HR in *N. tabacum* by the TIR-GFP-His6 construct and its mutant variants. *(B)* ATR1-dependent induction of HR by the full-length RPP1-WsB, wild type and nls mutant variants.

Supplemental Table 1.

Primer name	Sequence (5'>3')
RPP1-WsB F	*cacc*ATGGGTTCTGCAATGAG
RPP1-WsB Nde1 R	TC***CATATG***AGCTCTCATCCC
RPP1-WsB Nde1 F	caccAGCT***CATATG***GATATGCTG
RPP1-WsB Spe1-HA R	CTAGACTATGCGTAGTCTGGTACGTCGTAGGGATA***ACTAG T***ACAGTGGTCGCAGCTA
RPP1-WsB TIR *Spe1* R	***ACTAGT***ATCCATATGAGCTCTCATCC
RPP1-WsB NBS *Spe1* R	TTA***ACTAGT***TATTGTATCGTCATTGAG
RPP1-WsB NBS F	caccTATCATTCTCACAAATGG
RPP1-WsB LRR F	*cacc*GATAGTAGGCGTTTTATAGGG
RPP1-WsB E72A F	TCTTGGTGCCTTGATGCATTAGCAGAAATCATG
RPP1-WsB E72A R	CATGATTTCTGCTAATGCATCAAGGCACCAAGA
RPP1-WsB P-loop K293L F	CCGCCTGGGATTGGTCTGACAACTATCGCCAGA
RPP1-WsB P-loop K293L R	TCTGGCGATAGTTGTCAGACCAATCCCAGGCGG
ATR1Δ51-Emoy2/Maks9/Emco5 F	*cacc*ATGGCGCAGACAGCTC
ATR1Δ51 Emoy2/Maks9/Emco5 Spe1 R	TTA***ACTAGT***AACAGAATATATTCTCGAATACTCTTCCAC
ATR1Δ51-Cala2/Emwa1 F	*cacc*ATGGCGCAGGCAGCTCTCGATG
ATR1Δ51-Cala2/Emwa1 Spe1 R	CTA***ACTAGT***CTCGCTCGCGCGCC
ATR1-Maks9 E92K F	GATGAGGCTAAGATGAAGAAGGTCGTACGAACC
ATR1-Maks9 E92K R	GGTTCGTACGACCTTCTTCATCTTAGCCTCATC
ATR1-Maks9 D192G F	TCAAGCATTGCTACCGGCGACGATGTTGACTTT
ATR1-Maks9 D192G R	AAAGTCAACATCGTCGCCGGTAGCAATGCTTGA
ATR1-Maks9 D193H F	AGCATTGCTACCGACCACGATGTTGACTTTTTC
ATR1-Maks9 D193H R	GAAAAAGTCAACATCGTGGTCGGTAGCAATGCT
ATR1-Maks9 H183D F	CGAGGATACACCTCAGACAATGTCGTTTCAAGC
ATR1-Maks9 H183D R	GCTTGAAACGACATTGTCTGAGGTGTATCCTCG
ATR1-Emoy2 G192D F	TCAAGCATTGCTACCGACCACGATGTTGACTTT
ATR1-Emoy2 G192D R	AAAGTCAACATCGTGGTCGGTAGCAATGCTTGA
RPP1-NdA fragment 1 F	*cacc*GAATTCTCTTATTATATTTAAG
RPP1-NdA fragment 1 R	CATCCCAACAAAATCATCGAAATC
RPP1-NdA fragment 2 F	*cacc*ATGATTGAAAAGATATCCACTGAT
RPP1-NdA fragment 2 R	GATCCCAATCACTTTCACAA
RPP1-NdA fragment 3 F	caccCATTGATGTTCAAAGCTTGCTAT
RPP1-NdA fragment 3 R	GCCTATTTAGGCGGGCCTTACCTA
RPP1-NdA Xba1 F	GAATGCGGGATACTCCAA***TCTAGA***TAGATAGGAGACATT
RPP1-NdA Xba1 R	AATGTCTCCTATCTA***TCTAGA***TTGGAGTATCCCGCATTC

Supplemental Table 1. The list of primers used in this study for gene amplification and site directed mutagenesis. Restriction sites are shown in bold. The 5'CACC in the forward primers is required for introducing PCR products into pTOPO-ENTRY (Invitrogen) cloning system.

Supplemental Table 2.

dnaml		
U	Search for best tree?	Yes
T	Transition/transversion ratio:	2.0000
F	Use empirical base frequencies?	Yes
C	One category of sites?	Yes
R	Rate variation among sites?	constant rate
W	Sites weighted?	No
S	Speedier but rougher analysis?	Yes
G	Global rearrangements?	No
J	Randomize input order of sequences?	No. Use input order
O	Outgroup root?	No, use as outgroup species 1
M	Analyze multiple data sets?	No (for initial tree construction)
		Yes (for analyzing bootstrapped datasets)
I	Input sequences interleaved?	Yes
0	Terminal type (IBM PC, ANSI, none)?	ANSI
1	Print out the data at start of run	No
2	Print indications of progress of run	Yes
3	Print out tree	Yes
4	Write out trees onto tree file?	Yes
5	Reconstruct hypothetical sequences?	No
seqboot		
D	Sequence, Morph, Rest., Gene Freqs?	Molecular sequences
J	Bootstrap, Jackknife, Permute, Rewrite?	Bootstrap
%	Regular or altered sampling fraction?	regular
B	Block size for block-bootstrapping?	1 (regular bootstrap)
R	How many replicates?	1000
W	Read weights of characters?	No
C	Read categories of sites?	No
S	Write out data sets or just weights?	Data sets
I	Input sequences interleaved?	Yes
0	Terminal type (IBM PC, ANSI, none)?	ANSI
1	Print out the data at start of run	No
2	Print indications of progress of run	Yes
	Random number seed:	999
consense		
C	Consensus type (MRe, strict, MR, Ml):	Majority rule (extended)
O	Outgroup root:	No, use as outgroup species 1
R	Trees to be treated as Rooted:	No
T	Terminal type (IBM PC, ANSI, none):	ANSI
1	Print out the sets of species:	Yes
2	Print indications of progress of run:	Yes
3	Print out tree:	Yes
4	Write out trees onto tree file:	Yes

Supplemental Table 2. The set of parameters used in *dnaml*, *seqboot*, and *consense* algorithms for constructing the phylogenetic tree of the ATR1 alleles.

CHAPTER II: The *Hyaloperonospora arabidopsidis* ATR1 effector has distributed recognition surfaces and a structural subdomain conserved across oomycete species.

Abstract

The *in planta* association of the *Hyaloperonospora arabidopsidis* effector ATR1 with the cognate *Arabidopsis thaliana* RPP1 immune receptor activates a disease-resistance signaling pathway that inhibits pathogen growth. To define the molecular events specifying effector recognition by RPP1, we determined the crystal structure of ATR1 and assayed *in planta* the effects of surface polymorphisms that are critical to activating plant immunity. ATR1 adopts an all-helical, two-domain seahorse-like structure with an overall architecture unlike any previously described fold. Identification and mapping of critical recognition sites suggests that ATR1 detection by the RPP1 resistance protein is mediated by several distinct protein surfaces, which allow the effectors to escape recognition through natural surface polymorphisms. ATR1 gain-of-recognition mutants demonstrate that multiple amino acid substitutions are necessary for recognition and that surface polymorphisms exert additive effects. Structural comparison with another oomycete effector, Avr3a11, reveals a common hydrophobic core that creates a structural signature in a subset of oomycete effectors. These results suggest that ATR1 is a modular protein belonging to an ancient family of oomycete effectors that rapidly evolves to escape host detection and adopt diverse virulence functions.

Introduction

Oomycetes form a monophyletic group of organisms that morphologically resemble fungi but are evolutionarily more closely related to brown algae and *Alveolates* [78]. Oomycetes include a variety of commercially important plant pathogens with a diverse range of hosts, such as the infamous *Phytophthora infestans,* which causes tomato and potato late blight, *P. sojae* - stem and root rot of soybean, *P. ramorum* - sudden oak death, and *Plasmopara viicola* – grapevine downy mildew, as well as *Hyaloperonospora arabidopsidis (Hpa*, previously known as *Peronospora parasitica)*, a native pathogen of plant model organism *Arabidopsis thaliana*. Similar to many other plant and animal pathogens that deliver virulence effectors into the host to establish infection, oomycetes physically interact with their hosts through specialized haustorial feeding structures that facilitate the delivery of effector proteins into host cells where they have intracellular targets and play critical roles in oomycete survival and growth [19]. Despite substantial progress toward characterizing the roles of effectors, the unifying mechanisms by which oomycete effectors promote virulence remain largely unknown.

Oomycete effector genes have a number of conserved features. Although the mechanisms of effector translocation are not well understood, a typical eukaryotic signal sequence found in all effectors is thought to first mediate secretion out of the pathogen. Further translocation of effectors into the host cell is thought to be mediated by oomycete specific RxLR and dEER motifs [79-82], although it is not universally required [14]. Recently, genome sequences of *Hpa* and *Phytophthora* species allowed identification of numerous genes containing the conserved N-terminal signal peptide and RxLR-dEER motifs [11, 15]. A large family of *Phytophthora* effectors contains a conserved C-terminal W-motif [83], but only one has been identified in *Hpa* thus far [80].

The *Hpa/Arabidopsis* system has been adopted as a model system for studying plant-oomycete interactions. The *Hpa* genome encodes 135-150 putative RxLR-containing effectors [11, 15], including *Arabidopsis thaliana* Recognized 1 (ATR1) [45]. During the course of co-evolution, plants have developed surveillance systems dependent on highly polymorphic Resistance proteins (R-proteins). R-proteins directly or indirectly detect pathogen-derived effector molecules [43, 57, 84] to induce a cascade of immune responses that are collectively known as effector-triggered immunity (ETI) [4, 18]. The hallmark of ETI responses is the localized cell death, called the hypersensitive response (HR). Thus, pathogen effectors such as ATR1 have dual effects in promoting pathogen growth yet mediating recognition by the plant immune system through R-proteins.

Several dozen *Arabidopsis* R-genes have been shown to confer resistance to *Hpa,* including Recognition of *Peronospora parasitica* 1 (RPP1) [16, 44, 45]. RPP1, through its polymorphic Leucine Rich Repeat (LRR) domain, associates with ATR1 variants, leading to activation of plant disease resistance [16], This model of direct, physical interactions between ATR1 and RPP1 is supported by the apparent positive selection for sequence polymorphisms in both ATR1 and RPP1 [44, 45]. Moreover, recognition and

activation of immunity is mediated only by specific pairs of ATR1 and RPP1 alleles [45], suggesting a form of adaptive immune response in plants.

To explore the basis for ATR1 recognition by RPP1 and the mechanisms by which ATR1 mutations mediate escape from the host HR, we determined the crystal structure of ATR1 at 2.3-Å resolution. ATR1 is a monomeric, modular protein with two structural domains comprised of alpha-helices. Deletion analysis shows that two structural domains are required for recognition of ATR1 by RPP1. Naturally occurring polymorphisms of ATR1 allowed identification of surface residues critical for recognition. Multiple polymorphisms are needed to switch the specificity of ATR1 alleles, and the differential effects of mutations show that different alleles of RPP1 recognize distinct surfaces of ATR1. Comparison of ATR1 with the structure of the *Phytophthora capsici* Avr3a11 effector (Boutemy et al. 2011, in review) identifies a common structural subdomain in the region of ATR1 essential for RPP1 recognition and HR activation. Sequence alignment derived from this structural homology reveals that the W-motif previously identified in a large family of *Phytophthora* effectors forms a contiguous buried hydrophobic cluster that stabilizes a shared helical substructure in ATR1 and Avr3a11 [83]. This domain is conserved in at least 24 *Hpa* proteins. Or results show that ATR1 belongs to an ancient family of conserved oomycete effectors that evolves rapidly through surface polymorphisms to escape host recognition while maintaining a conserved structural core.

Results

ATR1 structure

Several ATR1 variants were expressed and purified from *E.coli*. ATR1Δ15, which included the RXLR/dEER translocation motif, expressed but was not amenable to crystallization. We focused our structural characterization on the effector domain sufficient for recognition by RPP1, ATR1Δ51 **(Figure 1a)**. We determined the crystal structure of ATR1Δ51 from *Hpa* isolate Emoy2 at 2.3-Å resolution (Rwork/Rfree=22.31/25.98), with initial phases generated by multiwavelength anomalous dispersion **(Table 1)**. Three copies of ATR1Δ51 crystallized in the asymmetric unit (AU). The three molecules in the AU are highly similar, with a Ca root-mean-square deviation (RMSD) of 0.43 Å. No clear electron density was obtained for the N-terminus (residues 51-62 of chain A, 51-62 of chain B, and 51-66 of chain C) and a loop connecting a12 and a13 (residues 278-290 of chain A, 280-289 of chain B, and 282-289 of chain C). 365 water molecules were built into well-defined density.

ATR1 adopts a two-domain, seahorse-like structure comprised of 13 α-helices **(Figure 1b)**. The N-terminal head (α1-α3) is separated from the larger C-terminal body (α4-α13) by a loop, or neck region (amino acid residues 117-126). The neck is stabilized by two β-turns, as well as several hydrogen bonds between Leu122 and Tyr126, Gly120 and His123, H123 and Thr125, and Asp124 and Asp127 **(Supplemental Figure 1a)**. Analysis of electrostatic surface potential of ATR1 reveals numerous electro-positive and -negative patches distributed across the surface, including a major

electropositive region on the head and two major electronegative regions on the body **(Supplemental Figure 1b)**. Mapping by sequence conservation among ATR1 alleles shows that polymorphic residues are on the surface of the head, neck, and body **(Supplemental Figure 2)**. Hydrophobic and aromatic patches occur on the exposed surfaces of helices a2 and a3 in the head domain, as well as a C-terminal pocket in the groove between helices a11, a12, and a13 containing six Phe or Tyr residues. Moreover, the first ordered residues in the structure, Trp-Pro-Phe 63-65, are unusually exposed for such hydrophobic amino acids.

Structural similarities to other proteins in Protein Data Bank

Comparison of the entire ATR1 effector domain with available structures using the DALI server did not reveal any significant structural homologs. Separate analysis of the N-terminal head and C-terminal body identified several potential distant homologs, with the circadian regulator KaiA (PDB ID 1R5Q) giving the best match (Z score = 5.8) to a four-helical segment **(Supplemental Figure 3a)**. ATR1 (α5-α8) aligns with a 4-helix domain of KaiA (residues 13-91) with an RMSD of 2.7 Å, but the electrostatic surfaces of the aligned structures are distinct. ATR1 also has a more extended loop between α5 and α6, and the KaiA helices analogous to ATR1 α7 and α8 are longer **(Supplemental Figure 3a)**. To test the putative role of ATR1 in clock regulation, we measured *Arabidopsis* circadian rhythms via the TOC1:LUC reporter in the absence and presence of ATR1. ATR1 had no effect on the transcriptional control of the circadian clock **(Supplemental Figure 3b)**.

ATR1 is a monomer *in vitro* and *in vivo*

ATR1 packed in the crystals as two equivalent dimers, with one formed by a crystallographic two-fold rotation axis. To determine if this dimer reflects solution properties of ATR1, we analyzed the oligomerization state of recombinant protein in solution by size exclusion chromatography (SEC). At 1mg/mL (28uM) in neutral pH buffer, ATR1 eluted at a volume corresponding to 30.2 kDa, similar to the molecular weight of a monomer, 29.5 kDa **(Supplemental Figure 4a)**. To test ATR1 stoichiometry in vivo, we performed co-immunprecipitations using FLAG- and HA-tagged ATR1 transiently expressed in *Nicotiana tabacum*. We used HA-tagged LRR of RPP1 as a positive control, as it has been shown to interact with ATR1 (10). FLAG-ATR1 was co-expressed with either HA-ATR1 or HA-LRR and immunoprecipitated using anti-FLAG M2 Sepharose. HA-ATR1 failed to co-immunoprecipitate with FLAG-ATR1, showing that ATR1 does not form homo-oligomers *in vivo* **(Supplemental Figure 4b)**.

Single amino acid changes in ATR1-Maks9 result in gain of recognition by RPP1-NdA

ATR1 alleles Emoy2 and Maks9 are different by five amino acids, yet only ATR1-Emoy2 and not ATR1-Maks9 is recognized by RPP1-NdA **(Figure 2a)**. We decided to

evaluate the contribution of those five amino acids to the pathogen's ability to escape recognition by RPP1-NdA. We used site directed mutagenesis to substitute each amino acid in ATR1-Maks9 to the corresponding residue in ATR1-Emoy2 and tested whether any of those single amino acid mutations could lead to recognition by RPP1-NdA. Two single substitutions E92K and D191G independently converted ATR1-Maks9 to be fully recognized by RPP1-NdA, while the other three sites did not have any visible effect **(Figure 2b)**. In the reciprocal experiment, when K92E or G191D mutations were introduced into the ATR1-Emoy2, either substitution delayed the recognition (the HR started to appear only at 48 hours post infiltration, data not shown), but did not reduce the intensity of HR once it started to appear. The double mutant K92E/G191D completely abolished recognition of ATR1-Emoy2 by RPP1-NdA **(Figure 2d)**. Western blot analysis showed that all of the ATR1 mutant variants produced protein amounts equal to the wild type **(Figure 2f)**. Additionally, the mutations did not have any effect on the recognition by RPP1-WsB **(Figure 2c, e)**, suggesting that additional sites can mediate interaction between ATR1 and RPP1-WsB. This genetic analysis suggests that it is unlikely that RPP1 monitors the enzymatic activity of ATR1, since several independent sites in ATR1 can activate recognition. It is interesting to note that the identified mutations involved charged residues, which can alter the charge of the surface area of ATR1 molecule, influencing its ability to associate with a cognate R-protein. Additional structural analyses will help us to understand the significance of those amino acid residues. This finding illustrates the narrow evolutionary line between recognition and susceptibility that puts effectors and *R*-genes under diversifying selection; in this case, a pathogen can be only one amino acid away from being recognized by the plant.

The minimal RPP1-WsB recognition region contains parts of both ATR1 domains

To elucidate the structural basis of ATR1 recognition by the host, we used deletion analysis of ATR1-Emoy2 to define the minimal region recognized by RPP1. We introduced deletion endpoints based on ATR1 secondary structure **(Figure 3a, b)** and assayed their activity by transient co-expression with RPP1-WsB in *Nicotiana tabacum*. The localized cell death due to the HR was used as a marker for activation of RPP1-mediated defense responses. Deletions of helices α1 and α2 (Δ67 and Δ87) from ATR1Δ51 did not affect activation of RPP1 but reduced ATR1 protein stability **(Figure 3d)**. Further N-terminal deletions failed to induce RPP1-dependent HR **(Figure 3c, d)**. Deletion of the C-terminal 90 amino acids compromised protein stability but did not affect recognition, suggesting that residues 87-222 are sufficient for RPP1 recognition. Further C-terminal deletions resulted in loss of HR **(Figure 3c, d)**, which may be due to lower protein stability or removal of critical amino acids. The minimal recognition region comprising amino acids 87-222 includes residues from both the ATR1 N-terminal head and C-terminal body **(Figure 3b)**.

Distinct ATR1 residues specify recognition by different RPP1 alleles

The natural polymorphisms between ATR1-Emoy2 and ATR1-Maks9 allowed identification of the two key residues that specify ATR1-dependent activation of RPP1-NdA. We extended this approach to define key amino acids that specify differential recognition of ATR1-Emoy2 and ATR1-Cala2 by RPP1-WsB. ATR1-Emoy2 and ATR1-Cala2 differ in 69 amino acids **(Table 2)**, 26 of which are located in the RPP1 recognition region (residues 87-222). Filtering for polymorphic sites that co-segregated between three ATR1 alleles recognized by RPP1-WsB (Emoy2, Maks9, Emco5) and two unrecognized alleles (Cala2 and Emwa1), yielding 16 sites that fulfilled these criteria **(Figure 4a)**.

We individually mutated these sites and assessed their relative contributions to activation of RPP1-NdA and RPP1-WsB in *N. tabacum* **(Supplemental Figure 5a)**. Substitutions at four sites produced gain-of-recognition phenotypes with RPP1-WsB that ranged from very mild (N158K), to intermediate (V122L, S125T), to strong (Y140D) **(Supplemental Figure 5a)**. Combining the mutations had additive effects, and the quadruple ATR1-Cala2 mutant (V122L/S125T/Y140D/N158K) induced HR with timing and intensity similar to wild-type ATR1-Emoy2 **(Figure 4b)**. Interestingly, activation of RPP1-NdA was not affected by any of these mutations. This reciprocal quadruple substitution in ATR1-Emoy2 significantly delayed activation of RPP1-WsB **(Supplemental Figure 5b)**, suggesting that although these four residues are sufficient to switch specificity, there are likely to be additional interaction sites. These ATR1 variants expressed to the same levels **(Supplemental Figure 5c)**, indicating that the changes in recognition specificity were not due to differences in protein stability.

Mapping these polymorphisms on the structure of ATR1 shows that they all are surface-exposed, except Asp140, which is partially exposed **(Figure 4c)**. Strikingly, these key recognition residues are not clustered in the structure. Instead, they are distributed across a large surface of ATR1. The residues important for activation of RPP1-WsB are chemically diverse—including positive, negative, and hydrophobic amino acids—and they are located on opposite sides of ATR1 in both structural domains. In contrast, amino acids in the neck region are important to activate RPP1-WsB **(Figure 4c)**.

Targeted mutagenesis in the LRR domain of RPP1-NdA expands its recognition specificity

Genetic manipulations of ATR1 guided by the natural polymorphisms allowed us to switch effector recognition. Next, we decided to test if the same result can be accomplished on the side of the host. Since successful recognition of ATR1 depends on its association with the LRR domain of RPP1, we used targeted mutagenesis of LRR domain of RPP1-NdA to broaden its recognition specificity. The LRR domain among many organisms has a conserved horse-shoe fold. Using structural homology modeling tools (PHYRE webserver), we constructed a structural model for the LRR of RPP1-NdA on the basis of solved structure of Human Toll / Interleukin Receptor 3 **(Figure 5a)**.

The LRR of RPP1 assumes a typical horseshoe shape composed of repetitive units and a concave side formed by parallel β-sheets. Previous reports show that LRRs tend to interact with their ligands through this concave side of the protein [85]. Overlaying natural polymorphisms between RPP1-NdA and RPP1-WsB on our structural model showed that 15 polymorphic sites map to the β-sheets on the concave side of the LRR. We mutated each of the polymorphic β-sheets in RPP1-NdA to the corresponding residues found in RPP1-WsB. Much to our excitement, one of the targeted regions, containing mutations D896K/S897Y/Y899R, allowed RPP1-NdA to recognize ATR1-Maks9, the effector allele that in wild type combination is not recognized **(Figure 5b)**. This activation of RPP1-NdA was effector-specific, not due to autoactivation (data not shown), and the mutated construct produced protein levels similar to wild type (data not shown). Another mutation in RPP1-NdA, D920H, produced a loss-of-recognition phenotype, **(Figure 5b)**, however, this protein variant was not detected by immunoblotting (data not shown), suggesting that it could have been destabilized. Collectively, our data provides first evidence that the recognition range of RPP1 can be altered by targeted manipulations of the LRR domain.

ATR1 shares a conserved sub-domain with the *Phytophthora* effector Avr3a11

Recently, the three-dimensional structure of the effector Avr3a11 from *Phytophthora capsici* was determined (Boutemy *et al.* 2011, in review), enabling comparisons with ATR1. Unexpectedly, Avr3a11 residues 68-132 structurally resemble a four-helix subdomain (residues 138-210) within the C-terminal body of ATR1 **(Figure 6a)**. This homology is undetectable by primary sequence comparisons (sequence identity=3.6%), but structural alignment between ATR1 and Avr3a11 reveals good agreement between the two regions (Ca RMSD = 2.8 Å) as well as conservation of several buried hydrophobic residues that make contacts with each other in the respective structures **(Figure 6b)**. The electrostatic surfaces in this region of ATR1 and Avr3a11 show considerable variation, despite similarity of overall fold **(Figure 6c)**. In contrast to Avr3a, this region of ATR1 occurs within the C-terminal body domain, and the conserved hydrophobic core is extended at each end through extensive contacts with the adjacent helices a4 and a9.

The conserved ATR1 core residues (Ile147, Val170, Trp174, Gly178, Tyr179, and Thr180) align structurally with the W-motif found in a large family of *Phytophthora* effectors [83]. To determine whether this structural homology reflects a shared virulence function, we tested if ATR1 inhibits PAMP-induced cell death, as previously shown for Avr3a. Our data show that, unlike Avr3a effector from *P. infestans*, ATR1 does not suppress INF-dependent cell death, suggesting a unique role for this subdomain **(Supplemental Figure 6)**. Identification of residues in ATR1 structurally analogous to the W-motif allowed us to refine a search for the W-motif containing subdomain among other *Hpa* effectors. Using structure-based amino-acid alignment between ATR1 and Avr3a11 **(Figure 6d)** and iterative Hidden Markov Model (HMM)-based searches, we identified 24 additional proteins in the *Hpa* genome that contain a putative W-motif containing subdomain (e-value < 1) **(Table 3)**. In several proteins,

this domain is present in 2 or 3 copies. A subset of the identified proteins also contained predicted secretion sequence (SignalP, HMM score > 0.75) and the effector specific dEER motif. Interestingly, one of the candidates (gene id 800198) had 3 tandem copies of the W-motif domain plus an N-terminal RxLR motif, but did not contain any detectible signal sequence **(Table 3)**. Overall, the consensus signature of the W-motif in *Hpa* **(Figure 6d)** is substantially different from the consensus among the *Phytophthora* proteins [83](Boutemy *et al.* 2011, in review).

Discussion

Although considerable progress has been made towards dissecting the molecular mechanisms underlying effector recognition by R-proteins and the structural basis for HR activation, virulence functions for many effectors remain elusive. Oomycete and fungal effectors have been shown to evolve under strong positive selection that drives rapid divergence, making it difficult to detect effector homologs outside their genus using primary sequence or secondary structure prediction tools. The three-dimensional structure of the *Hpa* effector ATR1 differs from that of other effector proteins, including AvrL567-A and AvrL567-D, two effectors from *Melampsora lini* (flax rust) differentially recognized by a cognate R-protein L [86], and *Pseudomonas syringae* effectors AvrPto and AvrB [87, 88]

ATR1 folds into an elongated structure composed of two major helical domains. The linker is stabilized by several hydrogen bonds that limit mobility. The similarity of all three independent copies of ATR1 in the crystals suggests that the linker may not function as a flexible hinge between the two domains. Although dimerization of R-proteins appears to be required in order to activate immune responses [16, 89]and several other fungal and bacterial effectors function as dimers [34]. ATR1 behaves as a monomer *in vivo* and *in vitro*. These results suggest that ATR1 does not serve as dimerization platform for RPP1. However, our results do not exclude the possibility that ATR1 may oligomerize in complex with RPP1 or other host partners.

Sequence-based searches for proteins with structural similarity to ATR1 did not yield significant matches, although several putative homologs were found by performing these analyses with individual domains. The highest scoring structural homolog was KaiA, a cyanobacteria *Anabaena* sp PCC7120 circadian clock protein [90] that aligned with ATR1 helices α4-α7. In light of this putative homology and the previously characterized link between circadian regulation and plant immune defense [91], we tested the role of ATR1 in circadian regulation. The presence of ATR1, however, does not significantly alter transcriptional regulation of *Arabidopsis* circadian rhythms. Given that this domain of KaiA is involved in promoting the kinase activity of its interacting partner KaiC, however, it is possible that the analogous region in ATR1 performs a similar biochemical function but a different physiological function. Other lower scoring homologs include an RNA-binding protein RBP8 (aligns with α4-α9) and Skp1 (aligns with α4-α10), an adaptor protein in the human SCF E3 ubiquitin-ligase complex. These structural similarities may offer initial insights into biochemical functions and potential partners of ATR1.

ATR1-Emoy2 recognition by RPP1-NdA relies on two polymorphic sites in *Hpa* strains--Lys92 and Gly191 [16]. These sites lie on opposite sides of the ATR1 structure in two different domains, suggesting that both positions can function in recognition. Unexpectedly, four natural ATR1 variations associated with recognition by another allele, RPP1-WsB, occur on a different surface of the protein. This gain-of-function phenotype shows that recognition of ATR1 can be achieved through different surfaces and supports the idea that RPP1 directly recognizes ATR1. LRR domains form the basis for adaptive immunity in jawless fishes [33]. The existence of two RPP1 alleles capable of recognizing different protein surfaces of ATR1 suggests that LRRs also play an adaptive role in plant immunity. Additionally, the additive effects of the ATR1 polymorphisms suggest that R-protein activation may be more complicated than a simple on/off switch.

Although unexpected from sequence comparisons, the structural homology of ATR1 α5-α8 to *P. capsici,* Avr3a11 defines and expands the role of the W-motif found in diverse *Phytophthora* effector proteins. The structural alignment of ATR1 and Avr3a11 shows that the W-motif comprises core hydrophobic residues that stabilize a shared four helical subdomain. This conserved core supports numerous surface variations that mediate escape from host recognition. This fold in ATR1 is not sufficient for suppressing INF1-induced cell death as seen with *P. infestans* Avr3a, which is consistent with the absence critical functional residues in ATR1 [92].

Interestingly, the W-motif sequence found in a family of *Phytophthora* effectors [83] differs from the structurally homologous W-motif in ATR1 and 24 additional proteins in *Hpa* (Figure 4C). This suggests that the W-motif is a structural signature that duplicated and elaborated in oomycete species after their divergence from the last common ancestor. The W-motif subdomain is appended to distinct domains in different proteins, suggesting that it functions as a modular, structural scaffold to enable functional diversification amongst effectors or allow rapid surface modifications to avoid host detection. Alternatively, this fold might have a specific biochemical function that is yet to be uncovered. Several W-motif proteins in the *Hpa* genome lack typical signatures of oomycete effectors, raising the possibility that this motif may have a more general function in oomycete biology.

In summary, our structure of the oomycete effector ATR1 reveals a novel, 13-helix, elongated fold. A linker connects two domains, one of which contains a diverged 4-helix subdomain shared by other oomycete effectors. Our mutational analysis reveals that RPP1 exhibits adaptive recognition of different ATR1 protein surfaces. Interpreting ATR1 polymorphisms in the context of ATR1 structure provides a framework for understanding how pathogens may escape detection and how plant hosts evolve in order to maintain effector recognition. Further understanding of the molecular mechanisms that allow R-proteins to respond to effectors could lead to engineering optimized plant pathogen receptors - potentially powerful new tools to contain some of the most important plant pathogens.

Materials and Methods

Strains and Growth Conditions

Bacterial DNA transformation was conducted using chemically competent *E. coli* DH5α (Invitrogen), electroporation of *E. coli* Rosetta (DE3) or through freeze/thaw transformation of $CaCl_2$ competent *A. tumefaciens* [76] . *Nicotiana tabacum* (variety Turk), *Nicotiana benthamiana* and *Arabidopsis thaliana* plants were grown in a controlled growth chamber at 24 ^oC at 16 hour light/ 8 hour dark photoperiod before infiltrations and switched to 24 hour light after infiltrations.

ATR1 cloning, protein expression, and purification

Previously described ATR1Δ51 Emoy2 deletion variant [16] was cloned into pDUET vector using BamHI/NotI restriction digest. The BamHI/NotI sites as well as the cleavage site for Tobacco Etch Virus (TEV) protease were added to the 5' and 3' ends of ATR1 through PCR amplification. For $(His)_6$-ATR1Δ51 protein expression, the pDUET construct was transformed into *E. coli* Rosetta (DE3) (Novagen), cells were grown at 37 ^oC to OD 0.7-0.8 and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside for 6-8 hours at 18^oC. The cells were harvested by centrifugation for 15 minutes at 5000 x g (4 ^oC) and frozen in liquid nitrogen for storage at -20 ^oC prior to purification. Cells were resuspended in Ni-A Buffer (20mM HEPES pH 7.5, 0.5M NaCl, 0.5mM TCEP, 25mM imidazole, 10% glycerol, 0.2mM AEBSF) and lysed by sonication. The lysate was centrifuged for 1 hr at 20000 x g (4 ^oC), and the supernatant was passed over a Ni affinity column. $(His)_6$-ATR1Δ51 was recovered by gradient elution with Ni-B Buffer (20mM HEPES pH 7.5, 0.5M NaCl, 0.5mM TCEP, 350mM imidazole, 10% glycerol, 0.2mM AEBSF) using the AKTA Explorer FPLC system. Fractions containing $(His)_6$-ATR1Δ51 were verified by SDS-PAGE and pooled for tag cleavage for 22 hrs at 4 ^oC with 1:50 TEV protease. Untagged ATR1Δ51 was loaded on the S75 gel filtration column and eluted as a monomer in 20mM HEPES pH 7.5, 100mM NaCl, 0.5mM TCEP. Protein was concentrated to 20 mg/mL.

Crystallization, structure determination, and structure analysis

Preliminary ATR1 crystals were obtained by hanging drop vapor diffusion trials at 18^oC from a 2:1 mixture of ATR1Δ51 at 20 mg/mL with 0.1M MES pH 6.5, 1.6M $MgSO_4$. Diffraction quality crystals were obtained by hanging drop vapor diffusion at 18^oC by seeding at 1:20,000 into a 4:1 mixture of ATR1Δ51 at 15mg/mL with 0.1M MES pH 5.0, 1.2M $MgSO_4$, 0.01% acetonitrile for 3-4 days. Crystals were dehydrated by transfer to a 0.1M MES pH 5.0, 1.5M $MgSO_4$ solution for 2 hours at 18^oC. Dehydrated crystals were immersed in mother liquor containing 14% ethylene glycol, mounted, and flash frozen in liquid N_2. Diffraction data were collected at the Lawrence Berkeley National Laboratory Advanced Light Source Beamline 8.3.1 [93]. Data reduction and initial maps were obtained using the automated ELVES program [94]. Phases were obtained

experimentally with data obtained from selenomethionine-substituted ATR1Δ51. The PHENIX software suite was used for initial model building. The final model was built by iterative manual model building using Coot and maximum likelihood refinement with PHENIX [95, 96]. Structure was validated using MOLProbity [97]. Images were generated in PyMol [98] and Chimera [99]. Multiple sequence alignment of the five ATR1 alleles was done using the MUSCLE algorithm [71] and visualized in CLC Genomics Workbench (www.clcbio.com). Structure analysis and comparisons were done using the DALI server [100] and Chimera [99] Coordinates and structure factors were deposited in the Protein Data Bank (ID code 3RMR).

Measuring the effect of ATR1 on circadian clock regulation in *Arabidopsis*

Transgenic Arabidopsis containing P(TOC1):LUC reporter (kindly provided by Frank Harmon, PGEC) were germinated and grown at 8/16 light dark cycle for the first three weeks, and then transferred into 12/12 light dark cycle conditions to train the circadian clock for additional 2 weeks. The night before the experiment these plants were sprayed with 5 mM luciferin, 0.01% Triton X-100. Next day, around noon, the leaves were infiltrated with *P. fluorescens* containing Type III delivery system and ATR1D49-Emoy2 in pEDV3 or pEDV3 empty vector. An additional $MgCl_2$ (no bacteria) control was included. Three hours later the inoculated leaves were detached, placed on Murashige and Skoog agar plates supplemented with carbenicillin (50 μg/mL) and sprayed with 5 mM luciferin, 0.01% Triton X-100 to boost luciferase activity. Bioluminescence rhythms were measured starting at 4 pm (ZT6) under constant light conditions. The measurements were taken every 2 hours. The data was analyzed using the Night Owl imaging system and BRASS software. Experiments were repeated at least three times.

Functional analysis of ATR1 in *Nicotiana tabacum*

The starting ATR1Δ51 pENTRY/TOPO constructs were cloned previously [16]. The deletion analysis was done through PCR amplification using the primers specified in Supplemental Table 1. The resulting products were introduced in pENTRY/TOPO (Invitrogen) and subsequently into pEG202 (35S promoter, N' FLAG tag fusion) [77] using LR clonase (Invitrogen). Site-directed mutagenesis was performed using the Quick-Change SDM Kit (Stratagene) with ATR1 pENTRY/TOPO template and subsequently introduced into pEG202. *Agrobacterium*-mediated transient expression in *N. tabacum* was performed as previously described [16]. Protein expression was sampled at 24-48 hours post induction and assayed as described previously [16].

Measuring the ability of ATR1 to suppress INF1-induced cell death

This assay was done essentially as previously described by Bos *et al. Plant Journal* (2006) **48**, 165-178. *A. tumefaciens* containing Avr3a, ATR1 or GFP control was infiltrated into *N. benthamiana* leaves at OD 600 = 0.3, and *A. tumefaciens* containing INF1 (kindly provided by Sophien Kamoun) was infiltrated at OD 600 = 0.3 around 24

hours later. The symptoms were observed at 3-8 days post infiltration and recorded at 4 days post infiltration.

Hidden Markov Model-based sequence searches

Structure-based amino acid alignment between the overlapping region in ATR1 and Avr3a was derived using PyMol and used as an initial seed in HMM building. The HMM building, calibration, searches and subsequent alignments were performed using HMMER software package, using *hmmbuild*, *hmmcalibrate*, *hmmsearch* and *hmmalign* respectively (hmmer.janelia.org). The HMMs were iteratively scanned for three rounds against the *Hpa* proteome (e-value cutoff of 1). Sequences were analyzed for eukaryotic signal sequences using SignalP 3.0 [101]. The W-motif domain sequence logo was made using WebLogo [102].

Figure 1.

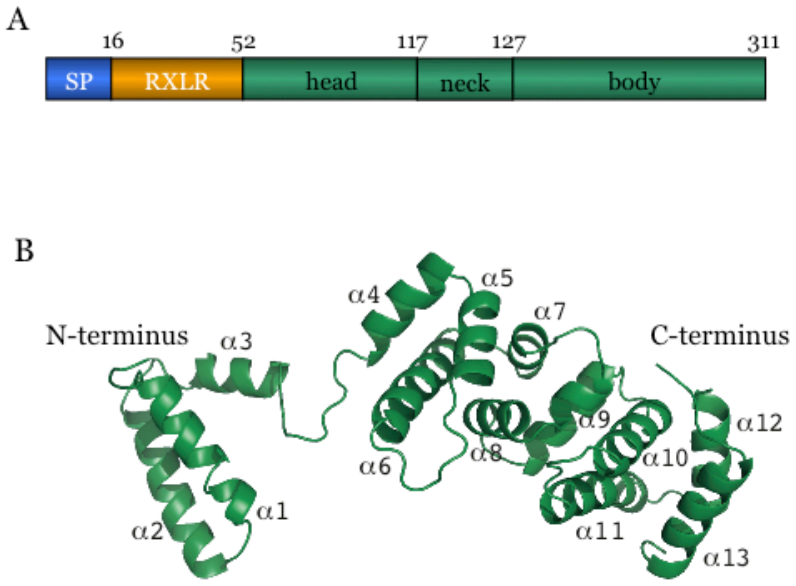


Figure 1. The structure of the RPP1-recognized domain of ATR1.
(A) A schematic representation of the domain architecture of ATR1. *(B)* Ribbon diagram showing the overall structure of the ATR1Δ51 that is bound by host Rpp1. Thirteen α-helices form N-terminal (α1-α3) and C-terminal (α4-α13) domains.

Figure 2.

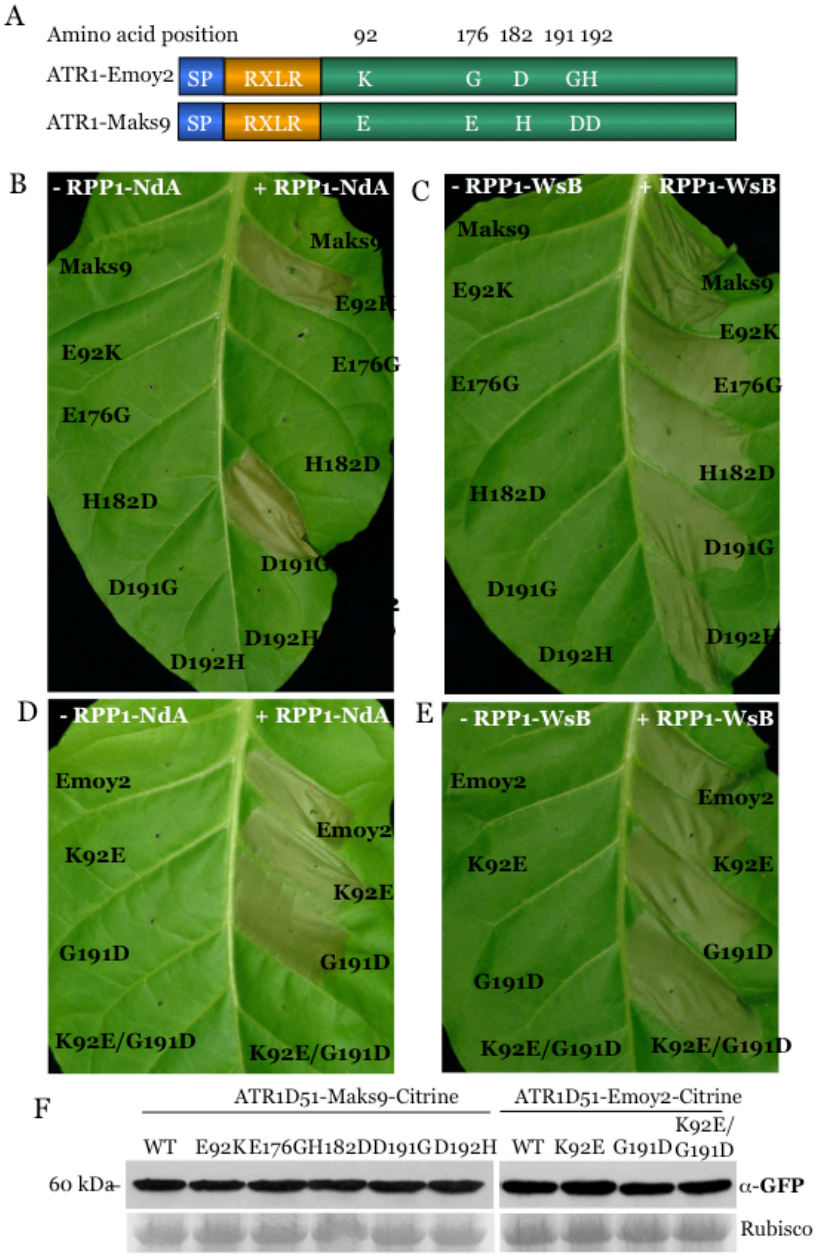


Figure 2. Two amino acids, Lys92 and Gly191, are critical for recognition of ATR1 by RPP1-NdA.

(A) ATR1-Emoy2 and ATR1-Maks9 protein sequences are different by five amino acid residues. B, C) Amino acid substitutions E92K or D191G convert ATR1-Maks9 into a variant recognized by RPP1-NdA, while not altering recognition by RPP1-WsB. D, E) K92E and G191D mutations in ATR1-Emoy2 collectively lead to loss of recognition by RPP1-NdA, but not RPP1-WsB. Pictures taken at 3 dpi. (F) Western blot showing similar levels of protein expression for all ATR1-Maks9 and ATR1-Emoy2 variants: WT - wild type. Rubisco – loading control.

Figure 3.

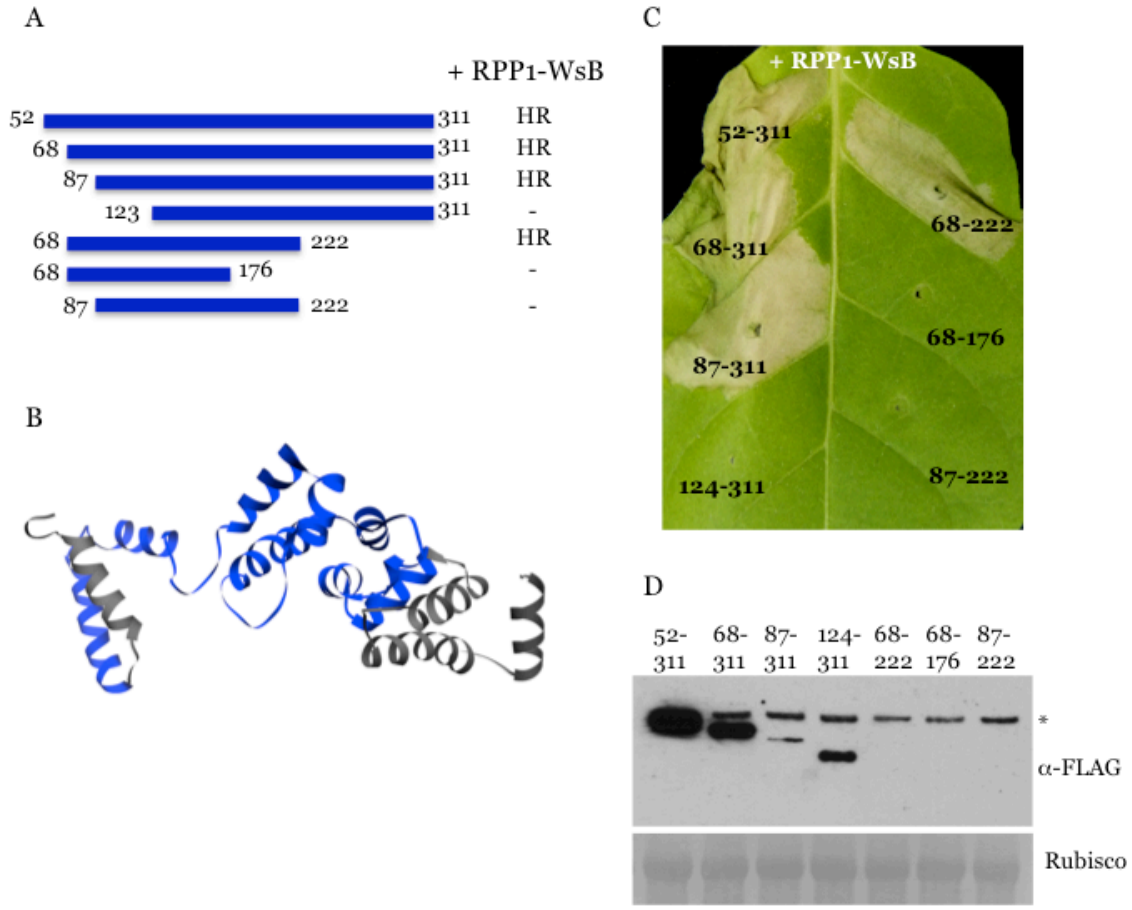


Figure 3. *In planta* deletion analysis of ATR1-Emoy2.
(*A)* Schematic diagram of the deletion constructs and summary of the hypersensitive response (HR) phenotype. *(B)* Ribbon diagram view of minimum recognition region (blue) of ATR1. *(C)* Co-inoculation of ATR1 truncations together with RPP1-WsB in *N. tabacum* showing the induction of RPP1-dependent HR. *(D)* Western blot showing relative protein levels of the truncated ATR1 variants. Asterisk marks a non-specific ~33-kDa band cross-hybridizing with α-FLAG antibody.

Figure 4.

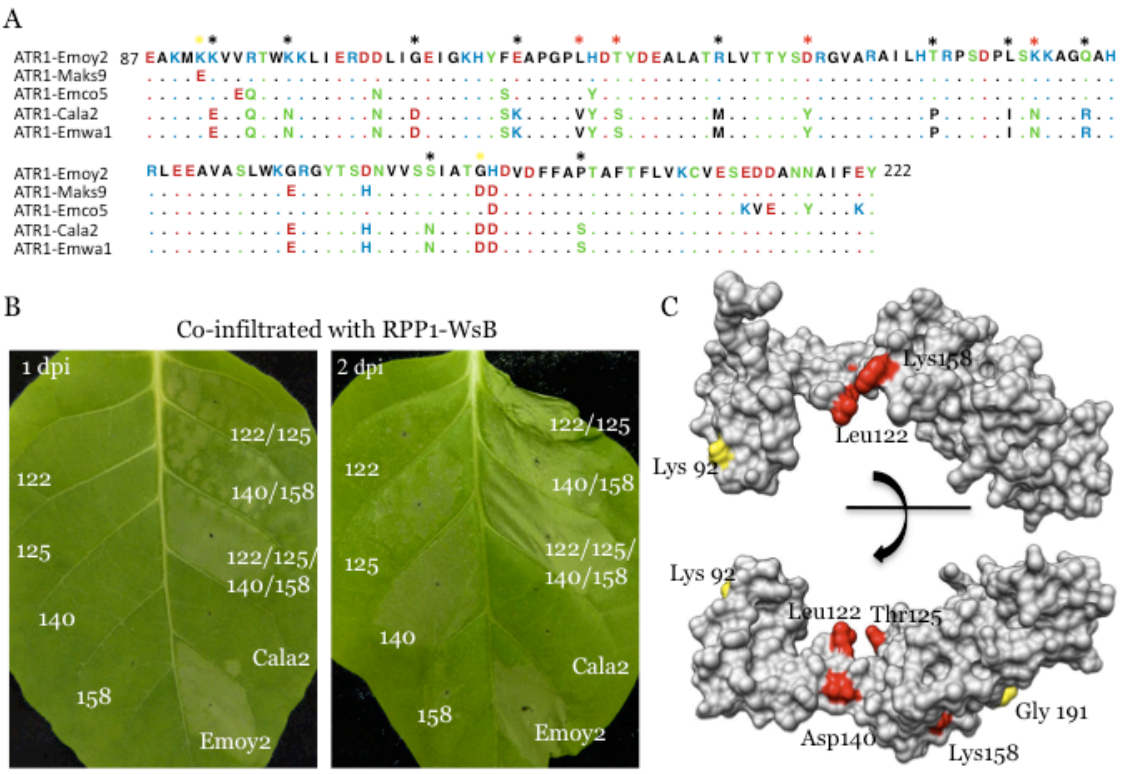


Figure 4. Several surfaces contribute to RPP1 activation by ATR1.
(A) Multiple sequence alignment of the 87-222 amino acid region of five ATR1 alleles. Previously identified amino acids critical for ATR1-Emoy2 recognition by RPP1-NdA are marked with orange asterisks. Residues that co-segregate with recognition of ATR1 by RPP1-WsB are shown with black asterisks. The four positions identified as critical for switching the specificity of HR activation are marked with red asterisks. *(B)* Additive effect of the four amino acids that contribute to recognition of ATR1 by RPP1-WsB as assayed by timing and intensity of HR in *N. tabacum*. The variants include: 122: ATR1-Cala2 V122L; 125: ATR1-Cala2 S125T, 140: ATR1-Cala2 Y140D; 158: ATR1-Cala2 N158K, and combinations of these substitutions, as well as a wild-type ATR1-Cala2 (negative control) and wild-type ATR1 Emoy2 (positive control).
(C) Surface representation of ATR1 with mapped polymorphisms, highlighting previously identified residues that were critical for recognition of ATR1 by RPP1-NdA (yellow) and the new mutagenesis data revealing that recognition of ATR1 by RPP1-WsB relies on different critical sites (red).

Figure 5.

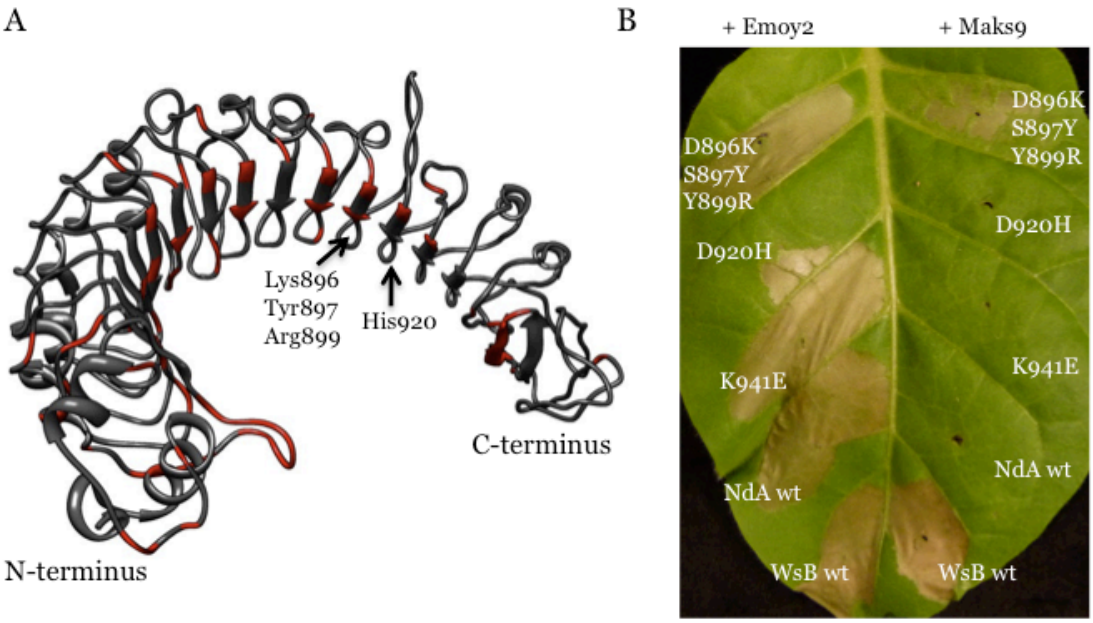


Figure 5. Gain of recognition mutation in the LRR domain of RPP1-NdA broaden its recognition specificity

(A) Structural model of the LRR region of RPP1 with polymorphisms between RPP1-NdA and RPP1-WsB shown in red. Arrows indicate two helices involved in changes of RPP1-NdA activity and corresponding critical amino acids. *(B) N. tabacum* leaf infiltrated with three RPP1-NdA mutants: gain of ATR1-Maks9 recognition (D896, S897Y, Y899R), loss of ATR1-Emoy2 recognition (D920H) and no effect (K941E). Wild type RPP1-NdA and RPP1-WsB controls are shown on the bottom of the leaf.

Figure 6.

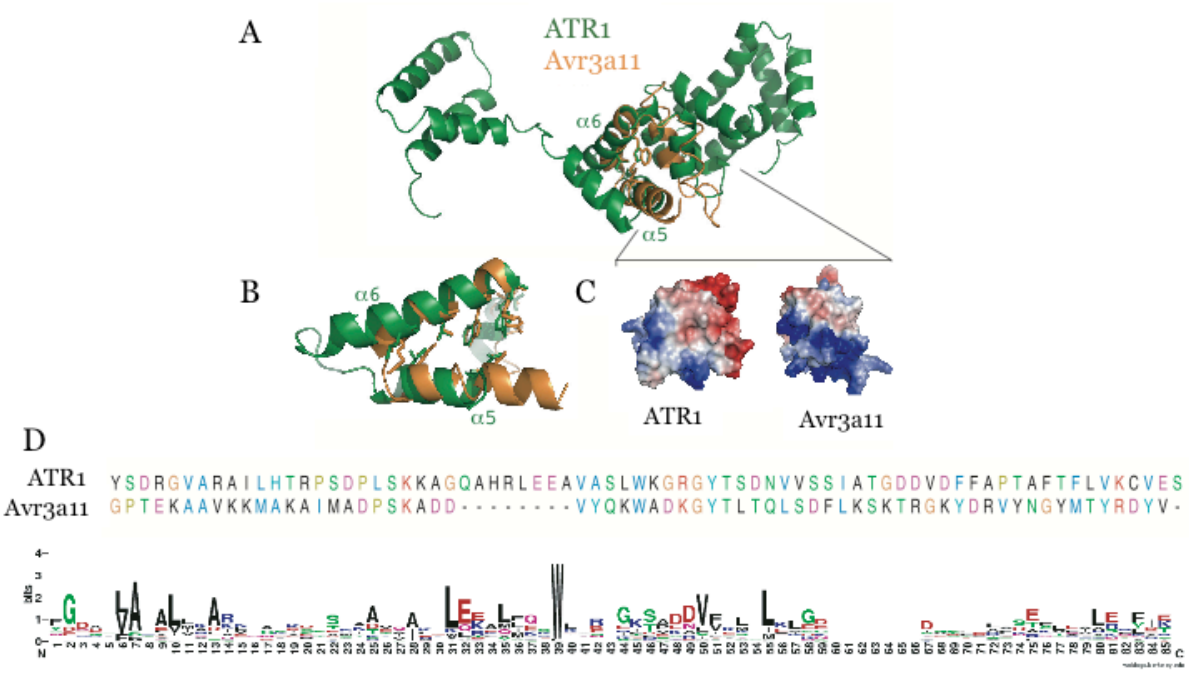


Figure. 6. Structural alignment of ATR1 and Avr3a11 reveals conserved core.
(A) ATR1 (green) and Avr3a11 (orange) share a common 4-helix domain in the body of ATR1. Sequence alignment of this region highlights conserved amino acids. *(B)* Close-up view of the conserved region. Structural-based sequence alignment shows a diverged the W-motif that forms a network of buried residues in the core. *(C)* Electrostatic surfaces of ATR1 α5-α9 and Avr3a11 shows diverged charge distributions displayed by the shared fold. *(D) Top* - Amino acid alignment between ATR1 and Avr3a11 derived from the structural comparison. Alignment is colored based on Clustal X coloring scheme. *Bottom* - Consensus sequence logo from the amino acid alignment of the W-motif in 25 genes identified from *Hpa*.

Table 1.

Crystal:	**Native:**	
Space group	P6122	
Cell dimensions	A=119.421, b=119.421, c=312.796	
Angles	90,90,120	
Dmin, Å	2.3	
No. of measurements	57,939	
No. of unique reflections	7,685	
% Completeness	97.5273	
I/σ	4.375	
R_{merge}(%)	8.3	
Copies in asymmetric unit	3	
Phasing (MAD):	**Fpeak:**	**Fhigh:**
Wavelength	1.0722	1.0631
Resolution, Å	2.84	2.84
Rsym	0.209(2.12)	0.205(2.00)
Completeness (%)	100	100
Multiplicity	30.6(30.3)	30.7(3.12)
I/σ	16.4(2.1)	17.7(2.3)
Phasing Power	0/0.26	0.03/0.25
Mean figure of merit (2.8-102.7 Å resolution)	0.099 (0.761 after solvent flattening)	
Refinement statistics:		
No. of reflections	55,573	
No. of atoms:		
Protein	5,482	
Ions	0	
Water	365	
Resolution, Å	62.3492-2.3	
Rcryst/Rfree	22.31-25.98	
Stereochemistry, RMS deviations:		
Bond length deviation, Å	1.02	
Bond angle deviation, ^{o}	0.015	
Average B factors, Å^2	47.2	
Ramachandran plot:		
Favored (%)	95.3	
Outliers (%)	0	

Table 1. Data collection and refinement statistics of ATR1.
Parentheses denote the highest resolution shell.

Table 2.

	Perc id.[a]	Differences[b]	NdA[c]	WsB[c]
Emoy2	100	0	+	+
Maks9	98.39	5	-	+
Emco5	85.03	47	-	+
Cala2	78.70	69	-	-
Emwa1	78.15	71	-	-

Table 2. Comparison of ATR1 alleles links amino acid identities to recognition by RPP1 alleles.

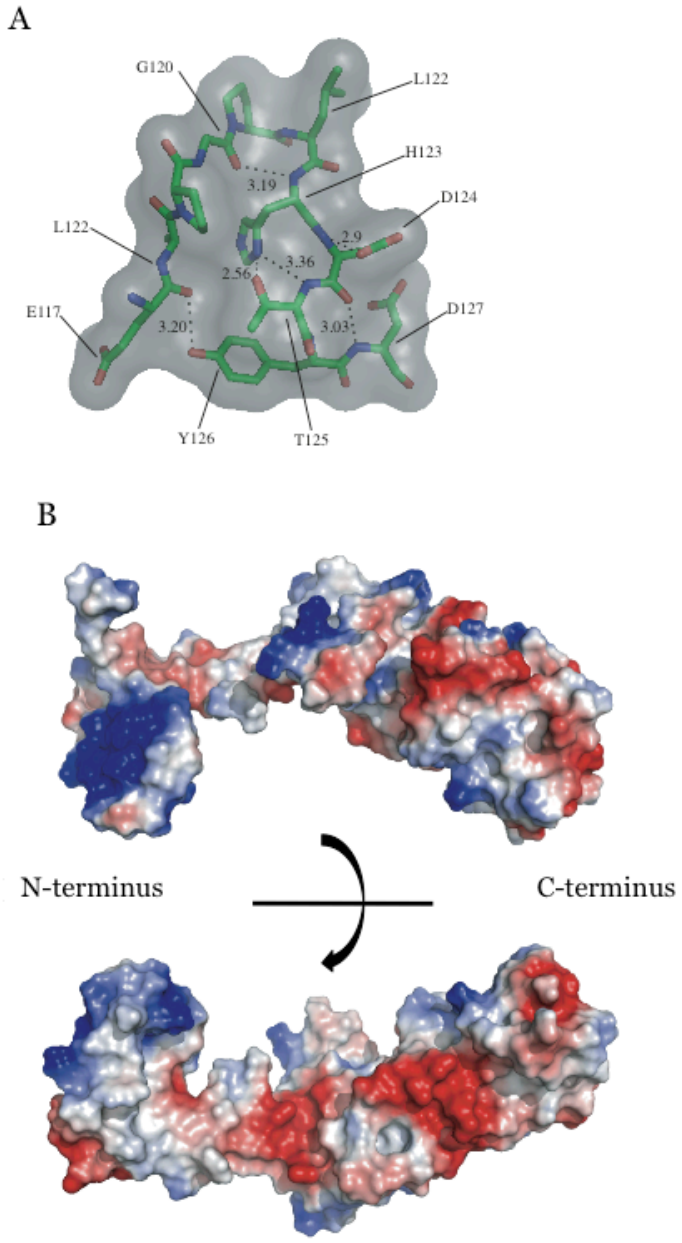
[a] Percent amino acid sequence identity to ATR1-Emoy2. [b] Number of amino acid polymorphism relative to ATR1-Emoy2, including gaps. [c] Recognition by RPP1 alleles NdA and WsB. + recognized, - not recognized.

Table 3.

Gene ID	Score	E-value	Number of domains	Features of effectors
801846	131.7	3.40E-36	1	SP (prob = 0.993), dEER
813261	118.7	2.70E-32	2	
809859	117.4	7.00E-32	1	dEER
801867<-ATR1	112.1	2.60E-30	1	SP (prob = 0.988), RXLR, dEER
802236	111.3	4.70E-30	3	SP (prob = 0.802), dEER
811478	95.3	3.10E-25	1	SP (prob = 0.949), dEER
812377	92.2	2.70E-24	1	
814615	75.2	3.40E-19	1	
808367	48.9	2.80E-11	2	
814280	45	4.30E-10	1	
800198	17.7	0.008	3	RXLR
808581	15.5	0.014	2	
813431	8.9	0.089	1	
811521	7	0.15	1	
806967	5.7	0.21	1	
807782	5	0.26	1	
807781	5	0.26	1	
810794	4.1	0.33	1	
808368	3.5	0.39	1	
802347	3.1	0.44	1	SP (prob = 0.830)
812044	2.7	0.49	1	SP (prob = 0.716), dEER
811880	2	0.59	1	SP (prob = 1.000)
811884	1.8	0.63	1	
810698	0.6	0.86	1	SP (prob = 1.000), dEER
808349	0.2	0.97	1	
801471	0.2	0.97	1	SP (prob = 0.760)
Total 25				

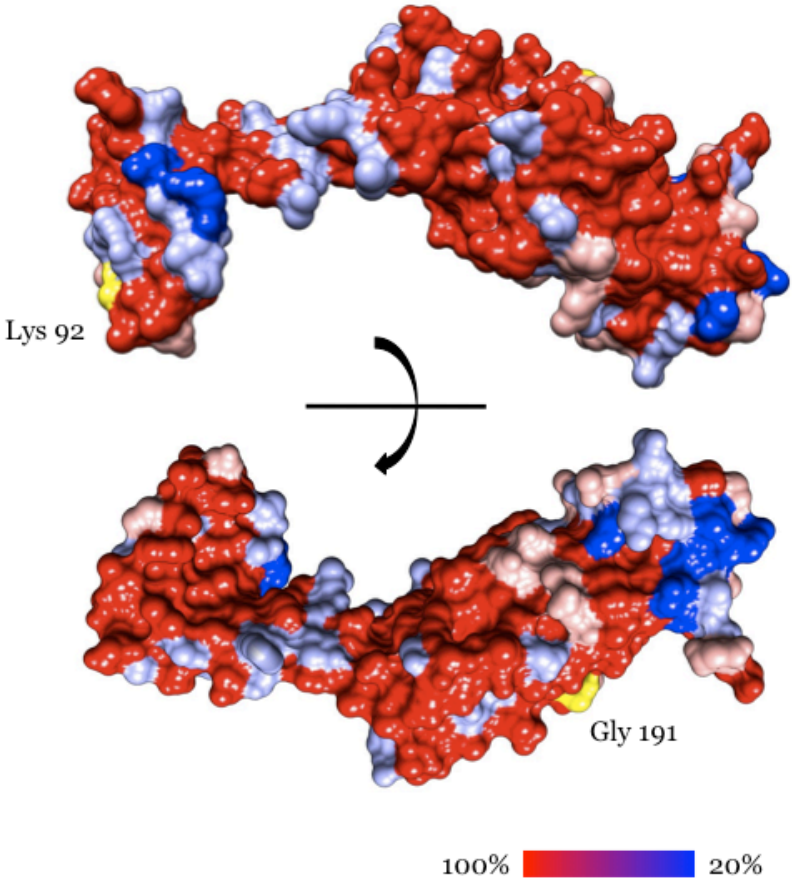
Table 3. Hidden-Markov-Model-based search for *Hpa* genes that contain a modified W-motif.

Supplemental Figure 1.



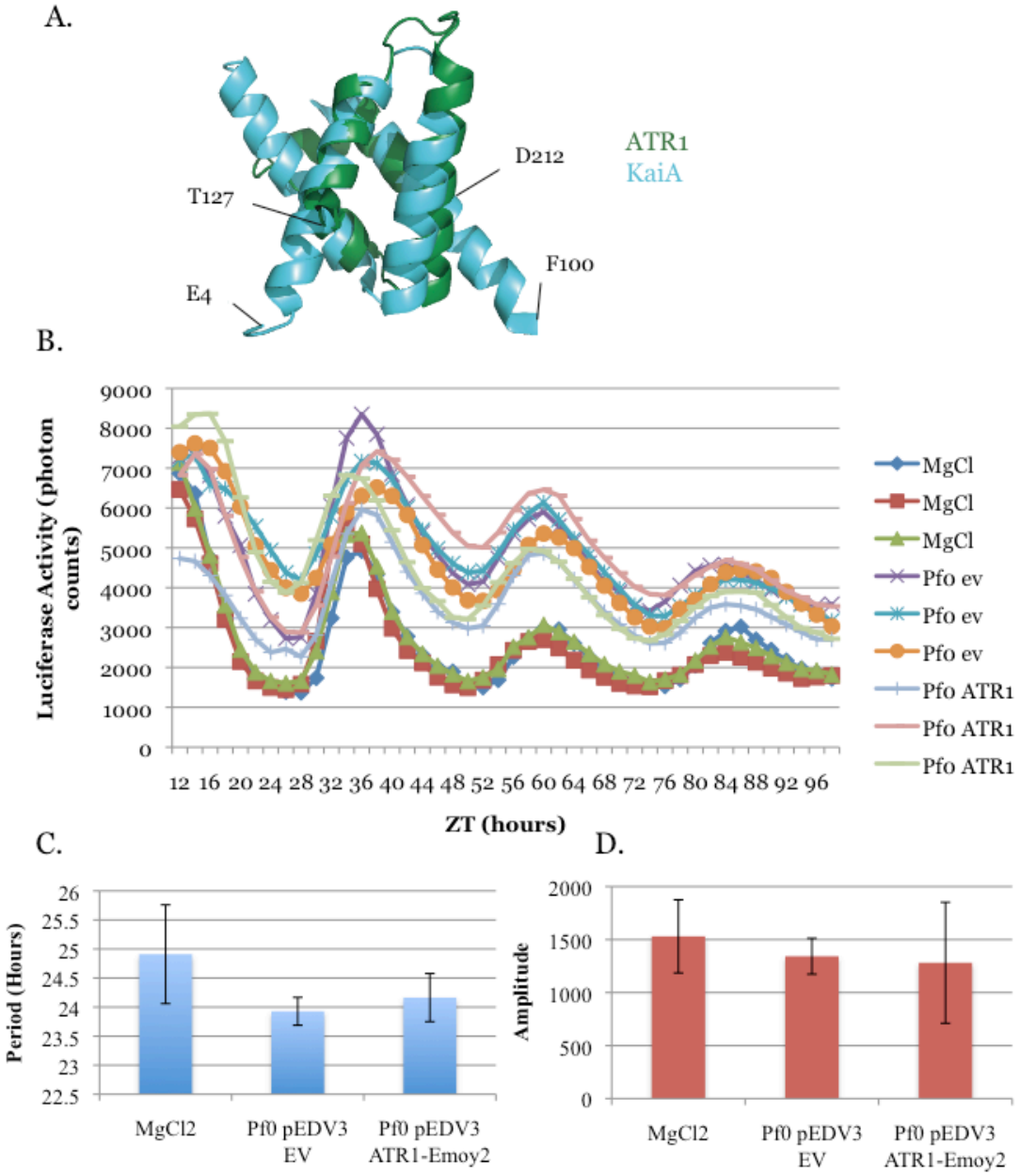
Supplemental Figure 1. Electrostatic Analysis of ATR1
(A) ATR1 neck is stabilized by hydrogen bonds. *(B)* A surface representation of ATR1 showing electropositive (blue) and electronegative (red) regions.

Supplemental Figure 2.



Supplemental Figure 2. ATR1 polymorphic residues are distributed across the surface of both domains.
Amino acid conservation levels between five ATR1 alleles (Emoy2, Maks9, Emco5, Cala2, and Emwa1) ranging from 100% conserved (red) to 20% conserved (blue) mapped on the structure of ATR1-Emoy2 using Chimera. In yellow, two sites polymorphic between ATR1-Emoy2 and ATR1-Maks9 implicated in recognition of ATR1 by RPP1-NdA.

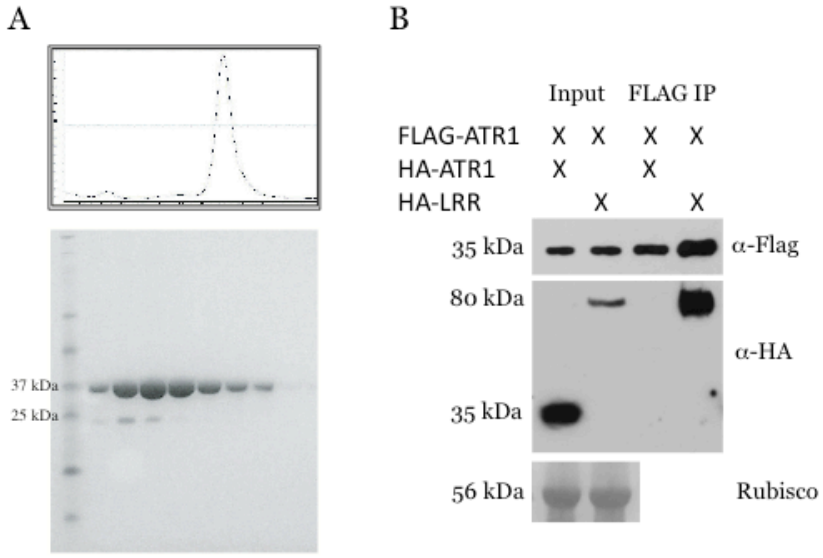
Supplemental Figure 3.



Supplemental Figure 3. ATR1 domain has structural homology to circadian regulator KaiA, but has no effect on plant circadian rythms

(A) Ribbon diagram of a structural alignment of ATR1 (green) and KaiA (cyan), generated by the *align* algorithm implemented in Pymol. The transcriptional output was measured using Arabidopsis Col-0 containing a stable P_{TOC1}:LUC transgene. ATR1-Emoy2 was delivered using the *Pseudomonas fluorescens* Type III delivery system and compared directly to an empty-vector control and $MgCl_2$ control (no bacteria). *(A)* Relative levels of luciferase activity measured over the period of 96 hours post infiltration. *(B)* Period and *(C)* Amplitude derived from the data shown above.

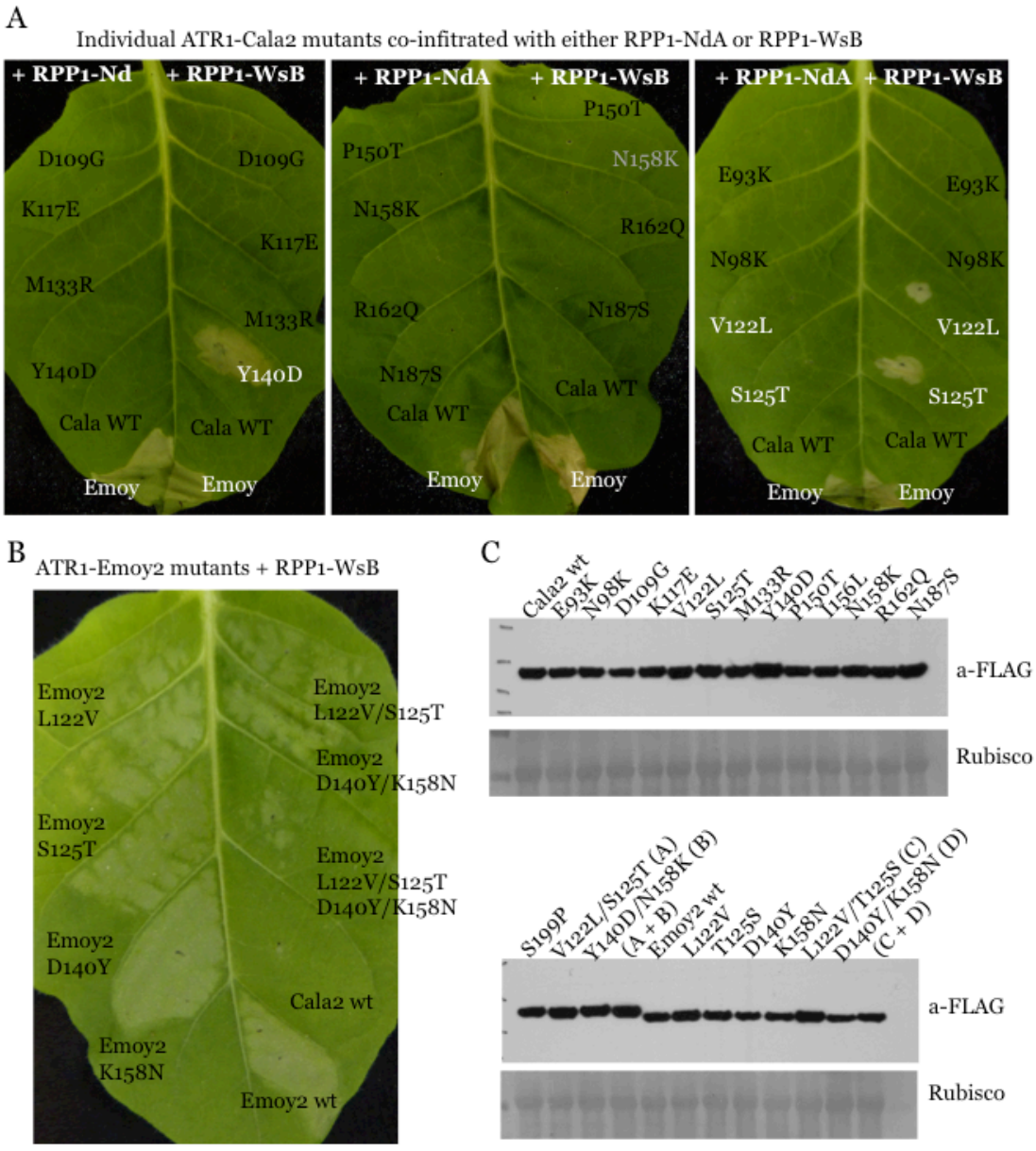
Supplemental Figure 4.



Supplemental Figure 4. ATR1 is a monomer in vitro and in vivo.
(A) Recombinant ATR1 elutes from S75 column at monomeric size *(B)* Co-immunoprecipitation experiment confirms that ATR1 is a monomer *in vivo*.

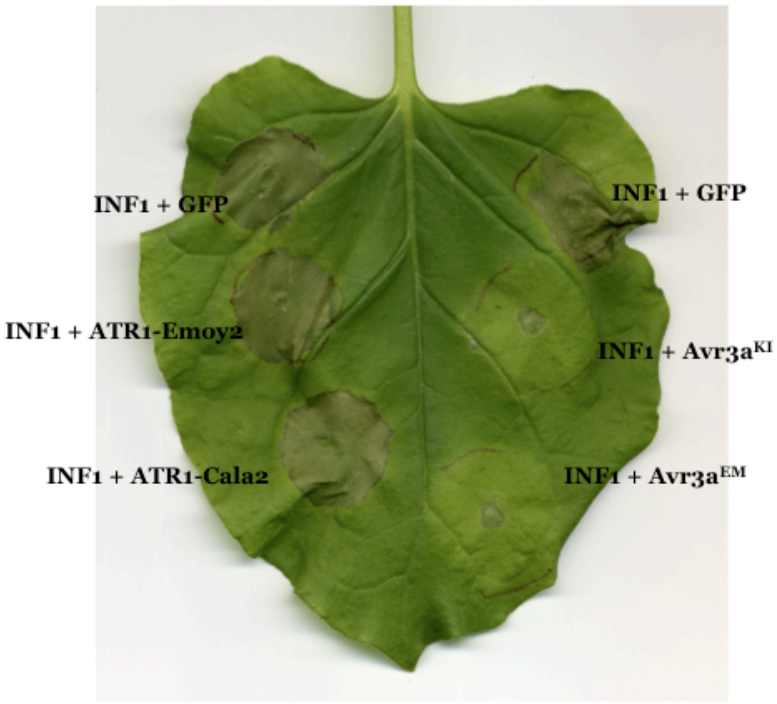
.

Supplemental Figure 5.



Supplemental Figure 5. Site-directed mutagenesis of ATR1.
(A) Phenotypes of the individual mutations introduced to ATR1 Cala2 and their relative contributions to recognition by RPP1-NdA (left side of the leaf) and RPP1-WsB (right side of the leaf). *(B)* Single and combined mutations in ATR1-Emoy2 showing their effect on recognition by RPP1-WsB *(C)* Relative protein expression levels of all ATR1 mutants used in this study. Rubisco – loading control.

Supplemental Figure 6.



Supplemental Figure 6. ATR1 does not inhibit INF1-induced cell death.
The necrosis inducing factor, INF1 co-infiltrated into *N. benthamiana* together with ATR1, with its known inhibitor Avr3a or GFP control.

Chapter III: Phenotypic variation among 83 *Arabidopsis* accessions reveals extreme levels of adaptation to its native downy mildew pathogen *Hyaloperonospora arabidopsidis*.

Abstract

Interactions between *Arabidopsis thaliana* and its native obligate oomycete pathogen *Hyaloperonospora arabidopsidis (Hpa)* present a model system to study evolution and natural variation of host/pathogen interactions. Both *Arabidopsis* and *Hpa* genomes are sequenced and collections of different sub-species are available. We analyzed interactions between 83 *Arabidopsis* ecotypes and 5 *Hpa* strains. We looked at the pathogen's sporulation, marking its ability to reproduce, as well as the amount of damage to the host due to pathogen growth and induction of defense responses. In parallel, we examined how much of *Arabidopsis* resistance can be attributed to recognition of ATR1 (*Arabidopsis thaliana* recognized 1), one of the *Hpa* effectors. Our results suggest that recognition of ATR1 is evolutionary dynamic and does not form a single clade in *Arabidopsis* phylogeny. We also demonstrate that strong developmental effect is prevalent among *Arabidopsis* populations, which is specified by yet undetermined factors. Finally, we show that the ultimate outcome of the interactions can be modified by the pathogen, despite functional gene-for-gene resistance in the host. Collectively, this data outlines different levels of regulation in *Hpa*/*Arabidopsis* interactions adapted by both the host and the pathogen that maintains the dynamic equilibrium of susceptibility and disease resistance.

Introduction

Hyaloperonospora arabidopsidis (*Hpa*, formerly known as *Peronospora parasitica*) is a native downy mildew pathogen of a plant model organism *Arabidopsis thaliana* [103]. *Hpa* is an obligate biotroph, growing only in life plant tissues; propagating to a new host by means of small asexual conidiospores that form on the plant leaf surface after the successful colonization of plant leaf tissues. Occasionally, sexual oospores form inside the plant, generating genetic diversity for the pathogen. In the case of host resistance, plant defense responses are induced shortly after the pathogen starts to grow. A visible hallmark of plant defense responses is induction of cell-death, also known as hypersensitive response. Genetic analyses of *Arabidopsis* disease resistance to *Hpa* have identified several dozens of disease resistance genes [44, 104-109]. Genetic and bioinformatic analyses in *Hpa* led to the identification of several confirmed effectors and prediction of 130-150 putative effector genes [11, 15]. The obligate nature of *Hpa*/*Arabidopsis* pathosystem put evolutionary pressure on both pathogen and the host. It has been shown that many *Hpa* effectors are under the pressure of strong positive selection [15, 38]. Similar evolutionary pattern is observed for the *Arabidopsis* disease resistance genes, often located in multiple copies at complex gene loci [38, 44]. Genetic and phenotypic diversity of *Hpa*/*Arabidopsis* interactions can give valuable insight into evolution and pathogenesis of obligate eukaryotic pathogens as well as plant immune system.

Current genomic projects aim at characterization of more than a thousand of *A. thaliana* sub-species, called ecotypes [110]. A set of 95 *Arabidopsis* ecotypes, collected from world-wide locations, and known as the Nordborg collection, has been characterized through small nucleotide polymorphisms and genome-wide association data is available [110]. Similarly, the genome sequence of *Hpa* stain Emoy2 is published [11], and several other *Hpa* strains, collected from the natural habitat are available [13]. The number of complete genomic sequences of strains and sub-species are rapidly accumulating due to development of high-throughput sequencing technologies. Taking full advantage of the genomic information requires careful phenotypic characterization of *Hpa*/*Arabidopsis* interactions.

There are two approaches to measure pathogen's interaction with the host: pathogen transmissibility, basic ability to complete the life cycle and send its progeny to a new host, and disease severity, amount of damage it causes to the host due to its activities or induction of host immune responses. Two previous studies addressed *Hpa/Arabidopsis* interactions on a population level. Eric Holub observed infected *Arabidopsis* cotyledons and developed an excellent descriptive scoring system, based on the amount and intensity of plant cell death, which was applied to a population of *Arabidopsis* collected in the United Kingdom [111]. A recent study analyzed the Nordborg collection and made macroscopic observations of infected true leaves, ranking them as susceptible, resistant or intermediate based on the presence of pathogen asexual spores [112]. However, a report observing interaction between *Arabidopsis* ecotype Col-0 and *Hpa* strain Emco5 showed that the interactions were controlled by host development; in this particular case, the pathogen was fully virulent on Arabidopsis

cotyledons, but failed to reproduce on true leaves [113]. Moreover, the amount of pathogen growth and plant cell death was substantially different on those organs [113]. Therefore, a more comprehensive study is needed to address the prevalence of this developmental control.

Much effort has been invested in investigating two known *Hpa* effectors, ATR1 and ATR13, on the molecular level. However, relative contributions of these two effectors to global *Hpa*/*Arabidopsis* interactions are not well understood. Contribution of bacterial effectors has been studied. It has been shown that recognition of conserved bacterial effectors is wide spread among *Arabidopsis* accessions and correlates well with the overall genomic variation between different *Arabidopsis* ecotypes [114], suggesting relatively slow rate of evolution of the cognate R-genes. On the other hand, oomycete effectors and corresponding *R-genes* show signatures of rapid evolution [15, 38, 45, 111], suggesting a different dynamics of interaction on a population level. A study of the *Arabidopsis* accessions from the United Kingdom shows that recognition of ATR13 can occur through two independent *Arabidopsis* loci [115]. Similarly, the RPP13 locus, originally identified to be responsible for recognition of ATR13, in some accessions can recognize a different oomycete effector [115]. Our recent studies on ATR1 suggest that its recognition in two *Arabidopsis* ecotypes could have evolved separately (see Chapter II). A study of oomycete effector recognition among different *Arabidopsis* accessions will advance our understanding of how plants co-evolve with oomycete pathogens to achieve recognition. Development of a surrogate oomycete effector delivery based on bacterial Type III Secretion System (TTSS) advanced our abilities to introduce individual effectors into the host. ATR1 and ATR13 delivered by TTSS induce resistance that is able to suppress growth of normally pathogenic bacteria in plants containing the cognate *R-genes*, *RPP1* and *RPP13* [39, 46]. Therefore, standard bacterial growth curves can be used as a quantitative measure for the resistance conferred by a particular *Hpa* effector. This allows to overcome challenges of working with an obligate, genetically untraceable *Hpa*, providing a rapid quantitative method to test how much of the observed disease resistance can be accounted for by the known *Hpa* effectors. Examining natural genetic variation in *Arabidopsis* will help us understand the evolution of plant disease resistance against highly specialized oomycete pathogens and guide towards new methods to engineer disease resistance.

In this study, we present a detailed analysis of *Hpa*/*Arabidopsis* interactions using a subset of ecotypes of the world-wide origins from the Nordborg collection and five *Hpa* strains, isolated in the United Kingdom. Examining each genotype-by-genotype interaction, we have recorded the ability of the pathogen to produce asexual spores, as well as amount of pathogen growth and extent of plant cell death. As a result, we developed a quantitative scoring system to describe five types of observed *Hpa*/*Arabidopsis* interactions. We have recorded our observations on both *Arabidopsis* cotyledons and true leaves, and observed a strong developmental effect of disease resistance in the wide population of ecotypes. Finally, we have used the TTSS delivery system to observe the prevalence of ATR1-dependend immunity among *Arabidopsis* ecotypes. Comparing plant response to ATR1 alone with the response to whole *Hpa* pathogen revealed that one of the *Hpa* strains, Emco5, can suppress recognition of ATR1. Our data addresses different levels of complexity in *Hpa*/*Arabidopsis*

interactions. We show that its outcome depends on several factors, including differential susceptibility of juvenile and adult tissues, as well as the pathogen's ability to evade recognition, although functional interactions between cognate effector and *R-genes* are in place.

Results

Resistance and susceptibility to *Hpa* is globally spread among *Arabidopsis* accessions

We have examined interactions between 83 accessions of *Arabidopsis thaliana*, collected from world-wide locations [116], and five strains of *Hyaloperonospora arabidopsidis* (*Hpa*) originally isolated in the United Kingdom [13]. The plants within a single ecotype did not exhibit substantial variation in response to any single *Hpa* strain. Overall, from 42% to 56% of all ecotypes supported asexual reproduction of at least one *Hpa* strain on cotyledons and from 27% to 50% – on true leaves **(Table 1)**. The *Hpa* strain Emco5 was least virulent on true leaves, producing asexual spores only on 27% of the examined ecotypes **(Table 1)**, similar to what have been previously reported [112]. However, our analyses suggest that *Hpa* Emco5 successfully colonized 42% of *Arabidopsis* cotyledons, comparable to other strains used in this study **(Table1)**. The overall pattern of disease resistance or disease susceptibility showed no clear correlation with geographic origin of *Arabidopsis* accessions. Similarly, there was no clear clustering of disease resistance based on the overall genome-wide phylogenetic relatedness of *Arabidopsis* ecotypes [116] **(Figure 1)**, suggesting complex evolutionary interactions between pathogen and the host.

Genotype-specific interactions between *Arabidopsis* and *Hpa* vary in amount of pathogen ingress and plant cell death

Asexual sporulation indicates the ability of *Hpa* to complete its life cycle and propagate, but it does not give a good measurement of the amount of pathogen growth or of the induction of plant immunity. Lactophenol trypan blue stain allows visualization of both, intercellular oomycete hyphae as well as the induction of the plant cell death [103]. To examine the relationship between host/pathogen interactions on the microscopic level and pathogen's ability to propagate, we performed lactophenol trypan blue staining of 83 *Arabidopsis* ecotypes inoculated with three *Hpa* strains: Emoy2, Emco5 and Emwa1 **(Appendix)**. Based on our observations, all *Hpa*/*Arabidopsis* interactions grouped into five cytological phenotypes, common to cotyledons and true leaves **(Figure 2a)**. We have ranked these phenotypes 1 through 5, ranging from resistant and less damaging to plant tissues to fully susceptible. The phenotypes are different from each other by two parameters: 1) the extent of pathogen growth and 2) the extent of plant cell death, which can either be radial, forming large circular patches of dying tissue (type 2 phenotype), or linear, tracing the pathogen hyphae (type 3 and 4) **(Figure 2a)**. We analyzed the correlation between microscopic phenotypes and ability of pathogen to

sporulate **(Figure 2b, Table 3)**. The Type 1 and Type 2 interactions successfully arrested *Hpa* growth and did not support any sexual or asexual sporulation. Type 3 phenotype, which showed intermediate levels of pathogen growth and some cell death supported sporulation in 55% percent of genotype-by-genotype interactions. Type 4 phenotype was marked by extensive pathogen growth coupled with plant cell death (commonly referred to as "trailing necrosis") and supported sporulation in 80% of interactions. Finally, Type 5 phenotype that lacks any signs of cell death was correlated with *Hpa* sporulation 100% of the time. This data clearly shows that ability of *Hpa* to reproduce is linked to its successful colonization of plant tissues, since it increases from phenotype 1 to 5. At the same time, the probability of supporting sporulation within each type of interaction is similar on both cotyledons and true leaves **(Figure 2b)**. On the side of the host, this data shows that its ability to induce cell death correlates with reduction in pathogen's ability to sporulate, but it is not enough to completely stop pathogen's growth. Since the "trailing necrosis" phenotype is correlated with reduced sporulation, it is likely to represent a hypersensitive response linked to a form of partially compromised plant immunity, rather than disease-related necrosis. We analyzed how many *Arabidopsis* accessions exhibited phenotypes 1 to 5 in cotyledons and true leaves in response to each *Hpa* strain **(Table 2)**. Phenotypes 1, 3 and 5 were most prevalent. Additionally, cotyledons were more prone to expansive plant cell death, manifested by phenotypes 2 and 4, which were nearly absent on true leaves **(Table 2)**.

The developmental resistance to *Hpa* is widespread among *Arabidopsis* ecotypes and follows gene-for-gene interactions.

Scoring the cotyledons and true leaves separately allowed us to quantify the prevalence of developmental resistance in the *Arabidopsis*/*Hpa* interactions. First, we compared whether pathogen could equally successfully propagate on cotyledons and true leaves. We observed that a substantial fraction of ecotypes, ranging from 4% to 12%, depending on the applied *Hpa* strain, consistently exhibited developmental effects, and more juvenile tissues, cotyledons, were more susceptible than true leaves **(Figure 3b)**. This developmental effect did not correlate with overall genome-wide relatedness between *Arabidopsis* ecotypes, nor with any particular *Hpa* strain, suggesting that it requires both plant and pathogen components **(Figure 1)**. Since presence or absence of sporulation can be variable on plants exhibiting phenotype types 3 and 4, we decided to compare the type of the phenotype scored on cotyledons with that on true leaves within the same genotype-by-genotype interaction. We observed that in 20% to 45% of all interactions, true leaves exhibited a different phenotype than cotyledons, and in 99% of these cases the extent of pathogen growth was higher on cotyledons than on true leaves **(Figure 3a)**. Since within the same type of microscopic interactions, *Hpa* has an equal chance to produce spores on cotyledons and true leaves **(Figure 2b)**, the resulting developmental effect should be due solely to the difference in the amount of pathogen growth. We make following conclusions from this data: i) different age plant tissues often respond differently to the same pathogen in terms of amount of pathogen growth and amount of plant cell death, ii) this developmental effect is globally spread among *Arabidopsis* ecotypes and does not correlate with overall genotypic relationship or

geography, and iii) it always follows gene-for-gene interactions, meaning that discrepancy in response between cotyledons and true leaves depends on both genotype of the plant and genotype of the pathogen.

Prevalence of ATR1 effector recognition among *Arabidopsis* ecotypes

Hpa, being an obligate pathogen, is currently not prone to genetic manipulations, therefore, making it difficult to assay relative contributions of individual effectors to either virulence or activation of disease resistance. Development of a surrogate bacterial delivery based on Type III Secretion System (TTSS) allowed studying *Hpa* effectors ATR1 and ATR13 in *Arabidopsis* [39, 46]. We have adopted a non-pathogenic *Pseudomonas fluorescens* (*Pfo*) supplemented with TTSS to minimize contribution of endogenous bacterial Type III effectors present in pathogenic *Pseudomonas*. This system allowed us to rapidly score recognition of ATR1 in all *Arabidopsis* ecotypes with minimal background. Delivery of ATR1 by *Pfo* into the known ecotypes containing cognate R-genes, *RPP1* (Nd-1 and Ws-0) and *RPP13* (Nd-1) induces a strong effector-dependent hypersensitive reaction (HR) at about 24 to 48 hours post inoculation **(Figure 4a)**. Using HR as our initial assay, we have screened *Arabidopsis* ecotypes with four polymorphic alleles of ATR1 (Emoy2, Maks9, Emco5, and Cala2). Our results show that there are four additional ecotypes that are able to recognize ATR1 **(Figure 4)**. Two of the ecotypes, Ws-2 and Pu2-23, have the same recognition specificity as Ws-0 being able to recognize ATR1-Emoy2, Maks9 and Emco5, but not Cala2. Another two ecotypes, Zdr-1 and Est-1, have altered recognition specificity, and recognize ATR1-Emoy2 and Maks9, but not ATR1-Emco5 or Cala2. The only ecotype able to specifically recognize ATR1-Emoy2 and not any other allele tested was Nd-1. Since plant cell death in *Arabidopsis* did not show 100% correlation with its resistance to *Hpa*, we performed bacterial growth curve assays delivering ATR1 through TTSS of *Pst* DC3000 **(Figure 4b)**. We observed perfect agreement between the HR induced in response to ATR1 delivered by *Pfo* **(Figure 4a)**, and restriction of *Pst* DC3000 growth **(Figure 4b)**. Unlike Sohn *et al.*, we did not observe any enhanced bacterial virulence in the presence of ATR1 **(Figure 4b)**. We compared evolution of ATR1 recognition with that of overall *Arabidopsis* genome. As overall disease resistance to *Hpa* did not show any obvious correlation with *Arabidopsis* geographic distribution or overall genetic relatedness **(Figure 1)**, the *Arabidopsis* accessions capable of recognizing ATR1 effector did not form a single evolutionary clade **(Figure 5)**. Moreover, groups of accessions with same recognition specificity towards different ATR1 alleles are more distantly related to each other than to those with altered recognition specificities **(Figure 5)**. This data shows that being the closest relatives with respect to overall genomes has little predictive power over the ability to recognize a specific oomycete effector. This could be due to more rapid rates of evolution in loci specifying disease resistance to *Hpa* compared to overall genome.

Hpa strain Emco5 can suppress ATR1-induced resistance.

Knowing the prevalence of ATR1 effector recognition allowed us to evaluate its relative contribution to disease resistance among *Arabidopsis* accessions. The contribution of ATR1 towards resistance varied depending on individual *Hpa* strain **(Table 4)**. Among 83 *Arabidopsis* accessions, 43 were resistant to *Hpa* Emoy2 and 6 of them recognized ATR1-Emoy2, thus recognition of ATR1 could contribute up to 14% of Arabidopsis resistance to Emoy2. A total of 36 ecotypes were resistant to *Hpa* Maks9, and recognition of ATR1-Maks9 was less prevalent (total of 5), therefore, overall contribution of ATR1 to resistance remained around 14%. Intriguingly, although resistance to *Hpa* Emco5 was prevalent (total of 48 ecotypes), recognition of ATR1 did not confer any resistance. All three ecotypes capable of ATR1 recognition were susceptible to *Hpa* **(Table 4, Figure 6)**. This effect is not due to artifacts of a bacterial delivery system since specific recognition of ATR1-Emco5 by Arabidopsis Ws-0 was also observed in biolistic bombardment assay [45] and by *Agrobacterium*-mediated transient expression [16]. ATR1-Emco5 gene was shown to be expressed in the pathogen [45], eliminating the possibility that this discrepancy was due to lack of gene transcription. Since *Hpa* is normally propagated at 18 ^{o}C and in high humidity, and bacterial assays are conducted at room temperature (around 20 ^{o}C), we addressed whether the discrepancy can be due to these conditions. We found no evidence for temperature or humidity regulation of ATR1 recognition, as the *Arabidopsis* plants were able to induce HR at 18 ^{o}C with same timing and intensity as at 20 ^{o}C or 25 ^{o}C and regular humidity levels (data not shown). The most likely conclusion from this data is the acquired ability of the pathogen, *Hpa* Emco5, to suppress recognition of ATR1.

Discussion

In the scope of this study, we provided phenotypic characterization of over 400 *Hpa*/*Arabidopsis* interactions. The interactions were analyzed from several different angles. As a result, we expanded our view of the system and the factors that control *Arabidopsis* disease resistance to *Hpa* on a global scale. These factors include i) wide prevalence of developmental control of disease resistance in *Arabidopsis*, ii) relatively small percent of resistance attributed to an individual *Hpa* effector, such as ATR1, iii) recognition of individual effectors, such as ATR1, is evolutionary dispersed and does not form a single clade on the level of whole genome phylogeny, iv) suppression of effector-triggered immunity by the pathogen.

Previously, our knowledge about developmental effects in *Arabidopsis* disease resistance to *Hpa* was limited to one isolated case [113]. Our results show that developmental resistance to *Hpa* is prevalent among *Arabidopsis* populations worldwide and follows gene-for-gene interactions. The effect is always directional with more juvenile organs, cotyledons, being more susceptible to the pathogen than true leaves. This effect is largely due to enhanced ability of the pathogen to colonize cotyledons and establish intercellular growth. The factors controlling this phenotypic difference between different plant organs is yet unknown. Since all of the known *R*-

genes that function against *Hpa* were cloned based on the resistance induced in cotyledons, we still do not know the primary source for the adult resistance of true leaves. Since our data shows that the developmental effect follows gene-for-gene interactions, it is unlikely that it is due to a mutation in a gene that has a global control of resistance pathways. The most parsimonious explanation is that a subset of yet unidentified *R-genes* is normally under developmental control and is only functional in true leaves. An alternative explanation can be postulated from the pathogen's perspective. In a subset of interactions, *Hpa* could be actively suppressing some of the resistance pathways in cotyledons. Both of those hypotheses imply that there is a difference in the disease resistance mechanisms in cotyledons and true leaves. A variety of plant phenotypes linked to phase change have been recently investigated and shown to be controlled by small RNA molecules [117]; it would be important to investigate whether they have a role in developmental regulation of plant immunity. Our data can be used to dissect the developmental effects through genetic crosses. Complemented with advanced sequencing technologies, it should be possible to map the source of developmental resistance in a variety of ecotypes.

Oomycete effector molecules, which serve as molecular triggers of plant defenses, form a class of extremely diverse and fast evolving proteins. These effectors alongside with plant R-proteins are molecular factors that specify dynamics of host/pathogen interactions on the evolutionary scale. Following individual effector/R-gene interactions, we can observe their contribution to the ultimate outcome of disease or resistance in a natural pathosystem. This knowledge can be translated to more sustainable control of more damaging agricultural pests. We looked at prevalence of ATR1 effector recognition among *Arabidopsis*, and found six ecotypes that recognized different subsets of ATR1 variants. Interestingly, these ecotypes did not form a single cluster on *Arabidopsis* phylogeny, suggesting that recognition of ATR1 could have evolved independently in different lineages. A similar conclusion is suggested by previous analyses of ATR13 recognition [115]. In the case of ATR13, it has been shown that its recognition can be specified by independent loci. Additionally, the same locus that specifies ATR13 recognition in some ecotypes can recognize a different effector in others. This shows *R-genes* that specify resistance against highly divergent oomycete effectors do not necessarily form families based on the effector that they recognize. Instead, a highly adaptive potential of a pool of genes provides genetic potential for maintaining effector recognition. This type of disease resistance, targeted at monitoring rapidly evolving molecules, is different from evolution of *Arabidopsis R-genes*, such as *RPM1*, *RPS2* or *RPS4*, that recognize effectors based on their enzymatic activity; the latter class of effectors is normally found under balancing selection [62].

The *Hpa*/*Arabidopsis* interactions have yet another level of complexity. This is due to the ability of the pathogen to escape host recognition without major modifications of effector gene sequence. The ATR1-Emco5 allele is recognized by several alleles of RPP1, recognition by RPP1-WsB has been previously demonstrated both in *Arabidopsis* [45] and by transient *Agrobacterium*-mediated expression in *Nicotiana tabacum* [16]. However, all of the ecotypes that are able to specifically recognize ATR1-Emco5 are susceptible to the *Hpa* Emco5 strain. This discrepancy cannot be attributed to genetic modifications of ATR1 and RPP1 coding sequences. There are several alternative

explanations for this pathogen escape of recognition. First, although it has been shown that ATR1-Emco5 is expressed in *Hpa*, we cannot exclude the possibility that it is not properly translocated inside the host. An alternative and more likely hypothesis is active suppression of ATR1 recognition or downstream signaling events by another *Hpa* effector. Suppression of effector-triggered immunity has been widely studied in case of bacterial effectors, but has yet to be demonstrated in *Hpa*/*Arabidopsis* interactions. ATR1-Emco5 interaction with RPP1 can serve as bait for uncovering immunity suppressors among the predicted *Hpa* effectors. Additionally, such suppression can introduce substantial noise to the genotype-based predictions about effector/*R-gene* interactions, and should be accounted for in evolutionary studies.

Together these studies open exciting new avenues for investigations of plant pathogen interactions. They point at ways to uncover the developmental regulation of plant immunity, providing a clear strategy of expanding a so far narrow pool of known ATR/RPP interactions. They also point at a possibility of active suppression of immunity by *Hpa*. Importantly, if non-allelic *R-genes* recognize the same effectors, and, on the other hand, allelic *R-genes* recognize different effectors, an update to the nomenclature of *R-genes* might be necessary to keep track of recognition specificities towards rapidly evolving effectors. Understanding the mechanisms controlling the dynamic equilibrium of host/pathogen interactions that is based on genetic diversity will allow development of more sustainable agricultural strategies, commonly relying on genetically restrained plant species.

Materials and Methods

Strains and growth conditions

Escherichia coli DH5a used for cloning and propagation of constructs was routinely grown at 37 ^{o}C in Luria Bertani broth media or agar plates supplemented with 10 μg/mL gentamycin. *Pseudomonas* strains were propagated at 28 ^{o}C. *Pseudomonas fluorescens (Pfo)* was grown on Pseudomonas Agar solid medium supplemented with 50 μg/mL tetracyclin, 30 μg/mL chloromphenicol and 150 μg/mL gentamycin and *Pseudomonas syringae* pv. tomato DC3000 (*Pst* DC3000) on NYGA solid medium supplemented with 100 μg/mL rifampicin and 5 μg/mL gentamycin.

Arabidopsis growth conditions, *Hyaloperonospora arabidopsidis* propagation and inoculations

The Nordborg collection of 95 *Arabidopsis* ecotypes, subset of which was used in this study, was described previously [116] and can be obtained from the Arabidopsis Biological Resource Center (ABRC, Ohio State University). A fraction of plants that routinely failed to germinate or had very delayed germination were dropped from the analysis, reducing the number of ecotypes from the original 95 to 83. For each experiment, a complete set of plants were grown in 2x2 inch pots and maintained at the same conditions, 24 ^{o}C growth chamber, 8/16 light-dark cycle. *Hpa* strains were

asexually propagated, and spray-inoculated on young *Arabidopsis* seedlings with the first set of true leaves. Conidiospore density in the inoculum was ~10^5 to 10^6 spores/mL, corresponding to roughly about 1-5 conidiospores per cotyledon or leaf. After inoculations, plants were transferred to the 18 ^{0}C chamber with high humidity. Inoculations were repeated at least three to four times. Sporangiophore formation was recorded at around 7-8 days post inoculation, when the oomycete life cycle has been completed. Trypan blue staining was done at 7-8 days post inoculation, following a previously described protocol [103] with minor modifications. Around 5-8 plants of each genotype were collected in 1.5 mL Eppendorf tubes with 0.5 mL of trypan blue stain. The tubes were boiled for 2 minutes and incubated on the bench for 2 hour to overnight. Seedlings were subsequently transferred to 96 well plates and de-stained overnight in 0.2 mL of chloral hydrate.

Type III delivery of ATR1, hypersensitive response assays and *Pseudomonas* growth curves

The ATR1Δ49-Emoy2 and Cala2 constructs cloned into the Type III delivery vector pEDV3 were kindly provided by Jonathan Jones (Sainsbury Labs, United Kingdom)[46]. The Maks9 and Emco5 alleles of ATR1D49 were subcloned into pEDV3 employing Sal1/BamH1 restriction enzyme cutting sites in the vector. All ATR1 constructs as well as empty vector pEDV3 were introduced from *E. coli* DH5a into *Pf*0 supplemented with Type III delivery system and into *Pst* DC3000 via triparental mating using *E. coli* HB101pRK600 helper strain. For plant inoculations, strains were grown from glycerol stocks on agar plates with appropriate antibiotics for 1-2 days. The hypersensitive response (HR) assays were conducted with *Pfo* inoculated at OD600 = 1.0 into young fully expanded leaves of 5-6 week old plants. Empty vector pEDV3 was included on each leaf as a negative control to monitor for any background plant response to *Pfo*. The HR was scored at 1-3 days post inoculation. Bacterial growth assays were conducted with *Pst* DC3000 using syringe hand-inoculation method as described previously [118]. Bacterial titer was determined at 0 and 3 days post inoculation.

Figure 1.

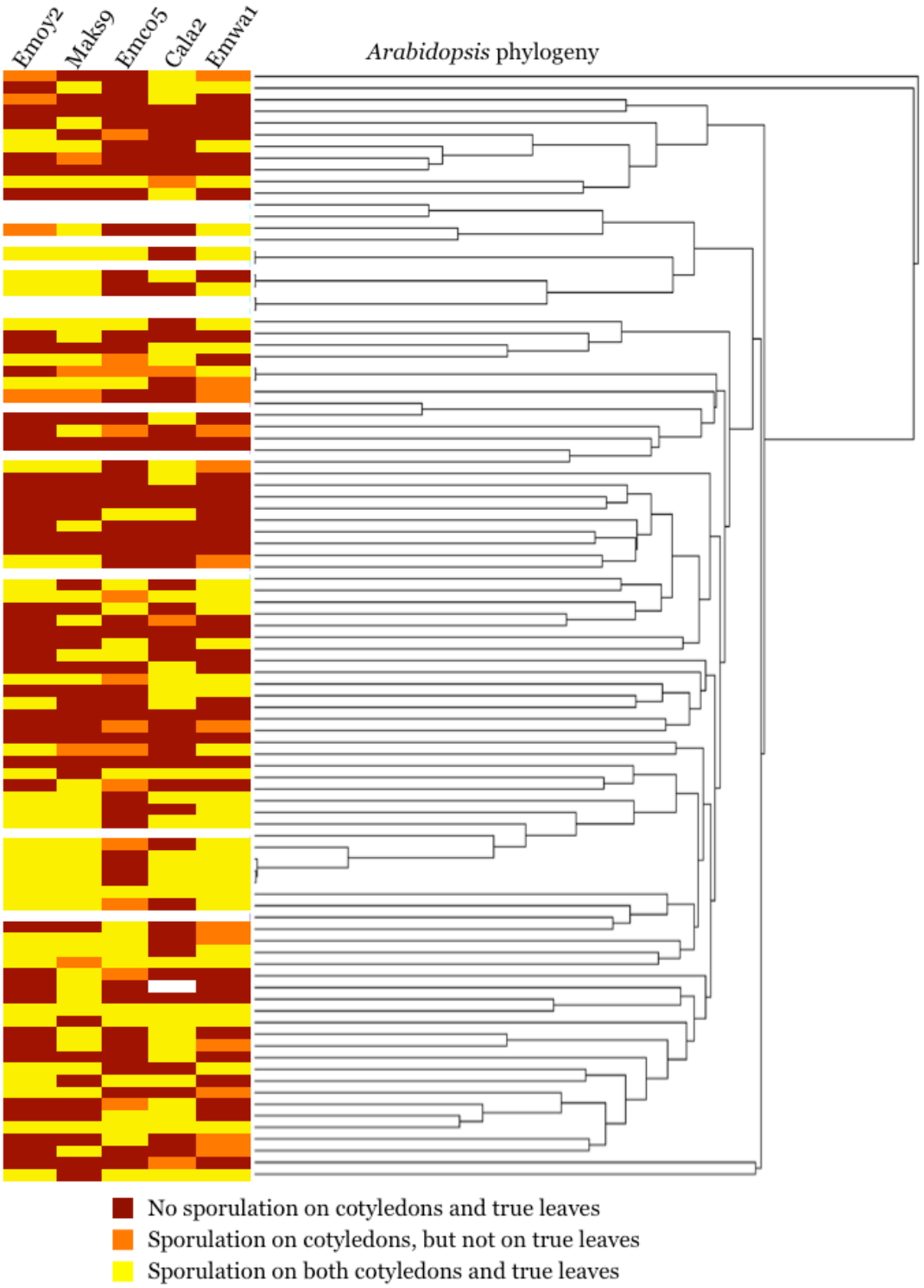


Figure 1. Resistance to *Hpa* compared to overall *Arabidopsis* phylogeny. The *Hpa* sporulation data obtained in this study is displayed on the left, tree on the right represents overall genome-wide relationship between Arabidopsis accessions [116].

Figure 2.

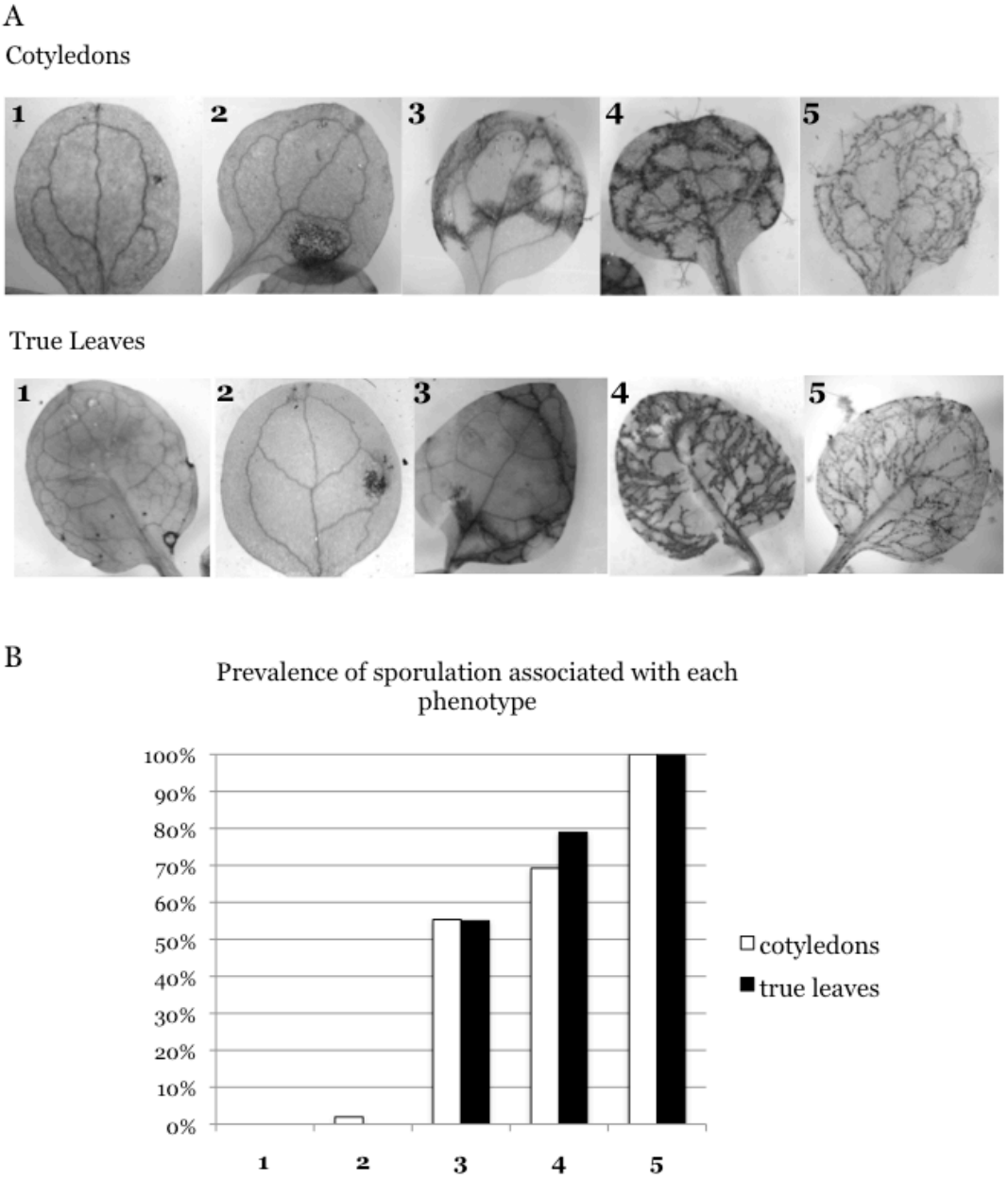


Figure 2. Five phenotypic categories defining race specific interactions between *Hpa* and *Arabidopsis*.
(A) Examples of the five phenotypic categories that were observed in cotyledons and true leaves. The following interactions are shown. Cotyledons: 1- Pu2-7/Maks9, 2 – Kz9/Emco5, 3 – Tamm-1/Emco5, 4 – Rmx-A180/Emoy2, 5 – Tsu-1/Emoy2. True leaves: 1 - Wa-1/Emoy2, 2 – Est1/Emwa1, 3 – Knox-18/Emoy2, 4 – Rmx-A180/Emoy2, 5 – Se-0/Emco5. *(B)* Prevalence of pathogen sporulation associated with each phenotype. Number of genotype-by-genotype interactions sampled, N = 270 for cotyledons, N = 301 for true leaves.

Figure 3.

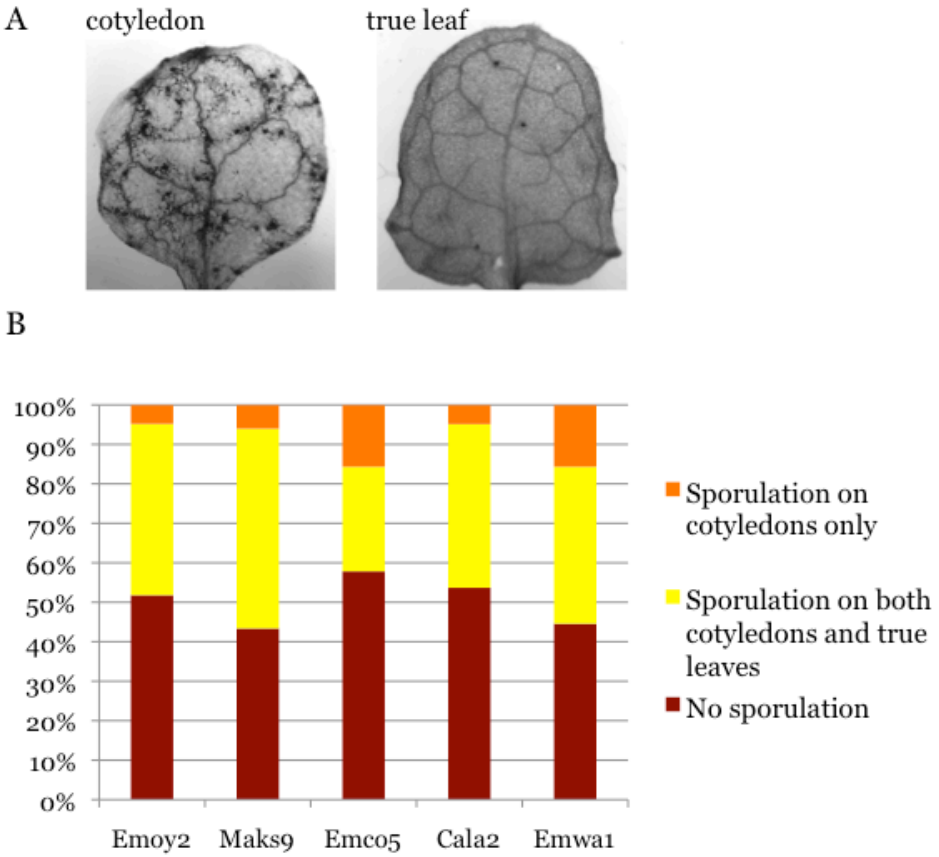


Figure 3. Arabidopsis cotyledons are more susceptible to *Hpa* than true leaves

(A) An example of *Arabidopsis* resistance to Hpa that shows developmental regulation. The interactions shown are between *A. thaliana* CIBC-5 / *Hpa* Emwa1.

(C) Prevalence of developmentally controlled resistance among the *Arabidopsis* ecotypes based on pathogen's ability to complete its life cycle. Number of ecotypes sampled, N = 83.

Figure 4.

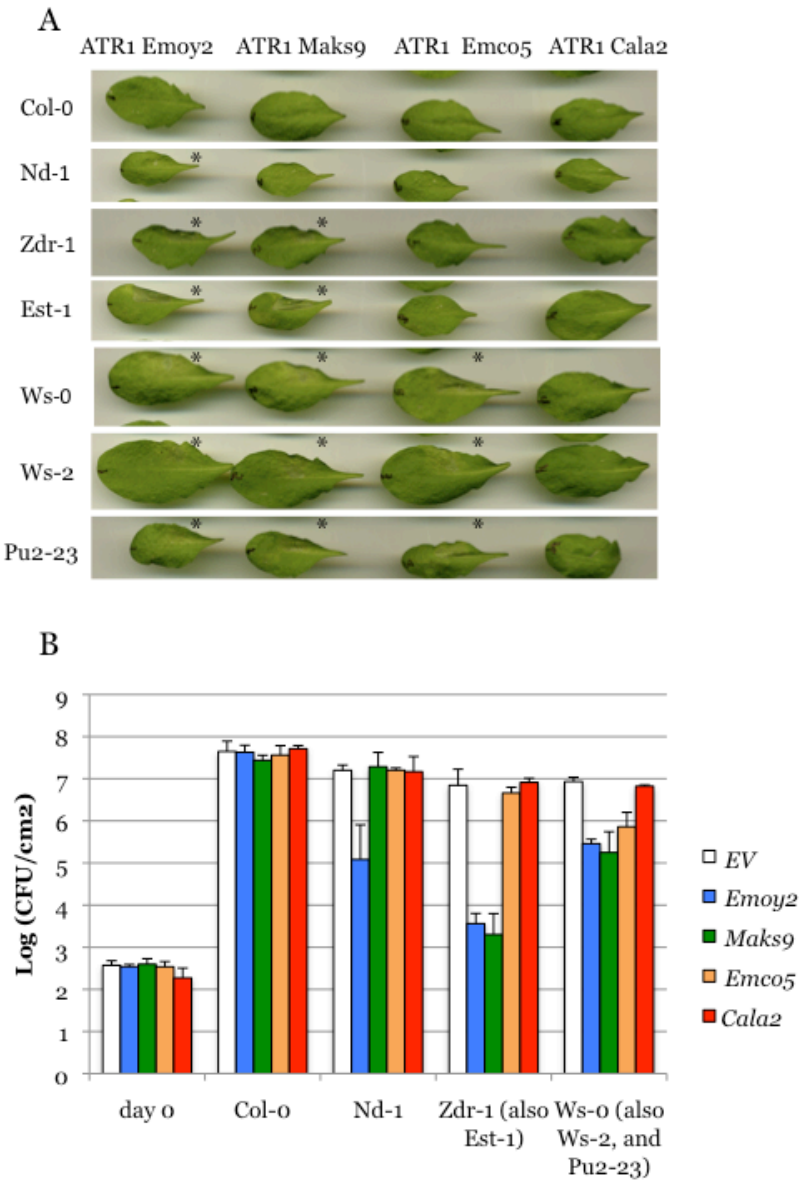


Figure 4. Six Arabidopsis ecotypes specifically recognize ATR1 effector.
(A) Recognition of ATR1 delivered by *P. fluorescens* induces HR in six Arabidopsis ecotypes. *P. fluorescens* carrying Type III delivery system and pEDV-3 empty vector, ATR1-Emoy2, ATR1-Maks9 , ATR1-Emco5 or ATR1-Cala2 infiltrated in Arabidopsis leaves and scored for HR two days post inoculation. The empty vector control was inoculated on each leaf (bottom left) alongside with ATR1 (top right).
(B) Representative growth curves show induction of ATR1-dependent resistance manifested by inhibition of bacterial growth. The same ecotypes as above were hand-infiltrated with *P. syringae* DC3000 carrying pEDV-3 empty vector, ATR1 Emoy2, ATR1-Maks9, ATR1-Emco5 or ATR1-Cala2

Figure 5.

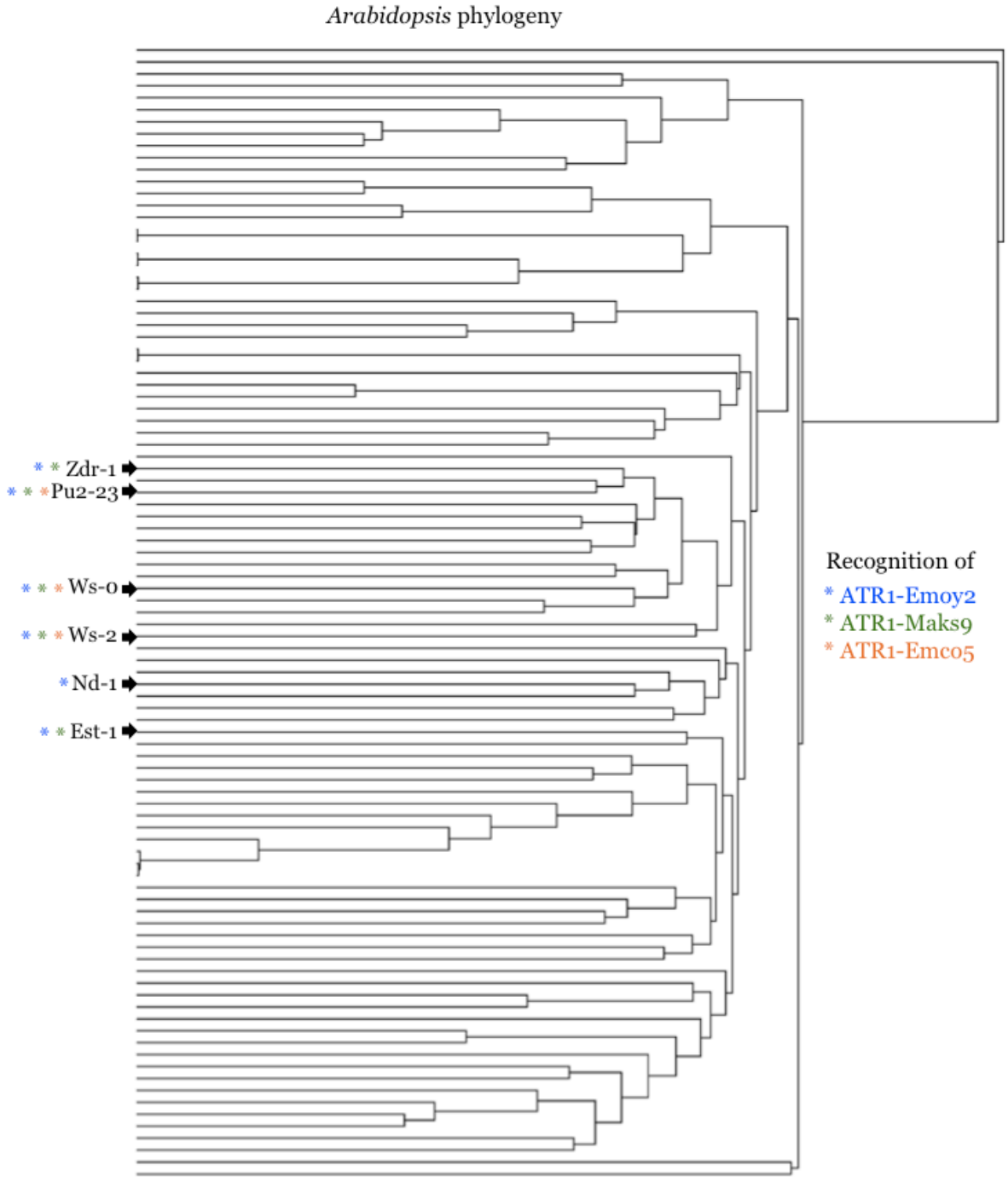


Figure 5. Phylogenetic relationship of ecotypes that recognize ATR1 effector relative to overall *Arabidopsis* phylogeny.
Six Arabidopsis ecotypes that are capable of recognizing subsets of ATR1 alleles are marked on the overall phylogeny of *Arabidopsis* ecotypes. The tree on the right represents genome-wide relationship between ecotypes [114, 116]. Asterisks mark subsets of ATR1 variants recognized by each ecotype.

Figure 6.

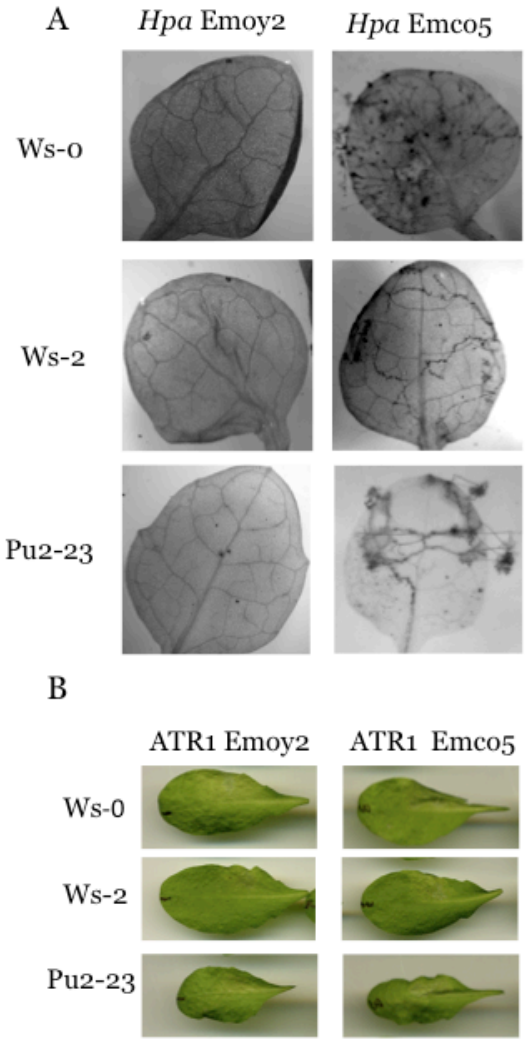


Figure 6. ATR1 effector recognition is actively suppressed in *Hpa* Emco5
(A) The ecotypes Ws-0, Ws-2 and Pu2-23 are able to induce defense responses to *Hpa* Emoy2, but not *Hpa* Emco5. *(B)* Both ATR1 alleles, Emoy2 and Emco5, delivered outside of the whole pathogen context are able to induce HR in Ws-0, Ws-2 and Pu2-23.

Table 1.

H. arabidopsidis strain

	Emoy2	Maks9	Emco5	Cala2	Emwa1
Cotyledons	48%	56%	42%	46%	55%
True Leaves	43%	50%	27%	41%	39%

Table 1. Percent of *Arabidopsis* ecotypes supporting *Hpa* sporulation
Total number of ecotypes inoculated with each strain, N = 83. Inoculations were repeated at least four times; ten to fifteen plants were examined in each experiment.

Table 2.

	Cotyledons				
	1	2	3	4	5
Emoy2	20%	25%	17%	4%	35%
Emco5	33%	17%	15%	6%	24%
Emwa1	16%	11%	27%	14%	32%

	True Leaves				
	1	2	3	4	5
Emoy2	42%	4%	17%	2%	36%
Emco5	56.5%	6%	14.5%	3%	20%
Emwa1	45%	6%	17%	7.5%	24%

Table 2. Percent of *Arabidopsis* ecotypes showing interaction phenotypes 1 to 5 on cotyledons and true leaves.
Interaction phenotypes were examined after lactophenol trypan blue staining of infected tissue. Total number of ecotypes, N = 83 total number of genotype-by-genotype interactions examined, 83*3 = 249. Five to ten plants of each ecotype were examined. The pictures used to score interactions are provided in the Appendix.

Table 3.

	Cotyledons				
	1	2	3	4	5
% sporulation	0%	2%	55%	69%	100%
Ratio S to N	0:64	1:50	31:25	18:8	104:0

	True Leaves				
	1	2	3	4	5
% sporulation	0%	0%	55%	79%	100%
Ratio S to N	0:114	0:14	22:18	15:4	83:0

Table 3. Association between sporulation and amount of pathogen growth / plant cell death (phenotypes 1 to5)
S – observed sporulation, N – no observed sporulation. Total number of different genotype-by-genotype interactions examined by lactophenol trypan blue included the 83 ecotypes sprayed with Emoy2, Emco5 and Emwa1, and partial microscopy data obtained for Maks9. The pictures used to score interactions are provided in the Appendix.

Table 4.

Response category		Number of ecotypes interacting with each *Hpa* strain/ATR1 allele				Explanation
Hpa[a]	ATR1[b]	Emoy2	Maks9	Emco5	Cala2	
S	No HR	40	47	32	38	No resistance.
R	No NR	37	31	48	44	Resistance is specified by other RPP/ATR interactions.
R	HR	6	5	0	0	Resistance is specified in part by ATR1/RPP1.
S	HR	0	0	3	0	RPP1 is functional, yet resistance is actively suppressed.

Table 4. Comparison between *Arabidopsis* response to *Hpa* and to ATR1 effector alone.

[a] inoculation with the whole pathogen, [b] delivery of ATR1 by Type III Secretion System. S: susceptibility (either cotyledons or true leaves), including sporulation. R: overall resistance (cotyledons and true leaves). No HR: no response, HR: hypersensitive response and inhibited bacterial growth.

Supplemental Table 1. Phenotypic responses of 83 *Arabidopsis* ecotypes to *H. arabidopsidis* strains Emoy2, Maks9, Emco5, Cala2 and Emwa1.

Accessions are listed in alphabetical order. Coloring scheme: brown – absence of asexual sporulation on both cotyledons and true leaves, orange – sporulation is present on cotyledons, but not on true leaves, yellow – sporulation is present on both cotyledons and true leaves. Numbers indicate phenotypic scoring (type 1 to 5, described in the text) for the interactions that have been analyzed by microscopy (see Appendix); first number in each column corresponds to a score on cotyledons second – on true leaves. S – sporulation, R – no sporulation.

		Response to *H. arabidopsidis* strain				
Accession	**Origin**	**Emoy2**	**Maks9**	**Emco5**	**Cala2**	**Emwa1**
Ag-0	France	2 n	5 5	1 1	n n	2 2
An-1	Belgium	2 1	5 4	5 n	R R	2 1
Bay-0	Germany	5 5	n 5	5 3	n R	5 3
Bor-1	Czech	2 1	2 2	1 1	R R	1 1
Bor-4	Czech	1 1	4 3	2 2	R R	1 1
Br-0	Czech	1 1	5 n	5 n	R R	3 1
Bur-0	Ireland	5 5	n n	5 5	S S	5 5
C24	Portugal	1 1	n n	2 1	S R	1 1
CIBC-5	England	3 1	5 5	1 1	R R	4 1
Col-0	Germany?	1 n	n n	5 n	R R	3 1
CS22491	Russia	2 n	n 4	2 1	R R	2 2
Ct-1	Italy	3 1	1 n	1 1	S S	1 n
Cvi-0	Cape Verde	2 2	n n	1 1	S S	5 5
Eden-1	N Sweden	3 1	5 4	4 1	R R	5 4
Edi-0	Scotland	2 1	n n	5 5	R R	3 1
Ei-2	Germany	5 5	n n	5 5	R R	5 5
Est-1	Estonia	1 1	1 n	1 1	R R	3 2
Fei-0	Portugal	5 5	n n	5 5	S S	4 4
Ga-0	Germany	4 5	3 3	5 2	S S	5 4
Got-22	Germany	3 5	4 5	5 5	R R	3 1
Got-7	Germany	2 1	4 n	5 n	S R	3 3
Gu-0	Germany	5 5	2 1	5 5	S S	5 5
Gy-0	France	3 3	5 4	2 n	R R	5 5
HR-10	England	2 1	2 n	5 3	S S	3 1
HR-5	England	2 3	2 1	4 3	R R	3 1
Kas-2	Kashmir	3 n	3 n	n n	R R	3 1
Kin-0	USA	2 n	n n	3 1	S S	2 2
Knox-10	USA	5 5	5 5	1 1	n S	5 5
Knox-18	USA	3 3	5 5	1 1	S S	4 3
Kondara	Tajikistan	3 3	4 4	1 1	R R	1 1
Kz-1	Kazakhstan	5 5	5 5	5 5	S R	5 5
Kz-9	Kazakhstan	1 1	2 2	2 1	S S	3 1
Ler-1	Poland	1 1	n n	1 n	S S	2 1
LL-0	Spain	5 3	n n	3 n	S S	1 1
Lov-1	N Sweden	5 5	5 5	2 1	S S	3 1
Lov-5	N Sweden	5 5	5 4	1 1	R R	3 3

Lp2-2	Czech	5 5	3 n	1 1	R R	3 3
Lz-0	France	2 n	n n	2 1	R R	1 1
Mr-0	Italy	3 1	n n	3 1	S S	3 1
Mrk-0	Germany	5 5	n n	5 4	S S	5 3
Ms-0	Russia	2 2	3 n	1 1	R R	2 1
Mt-0	Libya	1 1	3 1	2 1	R R	2 1
Mz-0	Germany	5 5	3 1	1 n	S S	1 1
Nd-1	Germany	2 1	n n	1 1	S S	5 5
NFA-10	England	5 5	5 5	1 1	R R	4 3
NFA-8	England	5 5	5 5	5 5	S S	4 4
Nok-3	Netherlands	2 n	n n	1 1	S S	4 3
Omo2-3	S Sweden	3 3	5 4	3 1	S S	3 1
Oy-0	Norway	5 5	n n	5 3	R R	5 5
Pna-10	USA	5 5	5 5	1 n	S S	5 5
Pna-17	USA	5 3	5 5	1 1	S S	5 5
Pro-0	Spain	5 5	4 n	5 5	S S	5 5
Pu2-23	Croatia	1 1	2 1	3 3	S S	1 1
Pu2-7	Croatia	1 1	1 n	1 1	R R	1 n
Ra-0	France	5 5	n n	5 5	S S	5 5
Ren-1	France	n n	5 5	1 n	S S	3 1
Ren-11	France	3 3	5 5	3 n	S S	3 1
Rmx-A02	USA	3 3	5 5	4 4	R R	5 5
Rmx-A180	USA	4 4	5 5	2 1	R R	5 5
RRS-10	USA	5 3	5 5	1 1	S S	5 5
RRS-7	USA	n 1	1 1	1 1	R R	n 3
Se-0	Spain	5 5	n 4	5 5	R R	4 4
Shahdara	Tajikistan	4 1	n 1	2 1	R R	1 1
Sorbo	Tajikistan	5 5	n 3	3 3	R R	5 5
Spr1-2	S Sweden	3 3	3 1	2 3	R R	4 3
Spr1-6	S Sweden	2 1	n n	1 1	R R	4 1
Sq-1	England	1 1	1 n	3 n	S S	1 1
Sq-8	England	5 5	5 4	2 1	R R	4 3
Tamm-2	Finland	5 5	n n	3 3	R R	5 5
Ts-1	Spain	2 1	n n	1 1	S n	5 3
Ts-5	Spain	2 n	4 n	2 2	R R	2 1
Tsu-1	Japan	5 5	5 5	5 5	S S	3 3
Ull2-3	S Sweden	5 5	5 5	3 3	S S	5 5
Ull2-5	S Sweden	1 1	4 4	4 1	R R	3 1
Uod-7	Czech	2 1	5 5	4 3	S R	2 1
Var2-6	S Sweden	1 1	2 1	1 1	n S	1 1
Wa-1	Poland	2 1	1 n	3 1	R R	3 1
Wei-0	Switzerland	5 n	n n	5 5	R R	5 5
Ws-0	Ukraine	1 1	n 1	5 5	R R	5 5
Ws-2	Ukraine	1 1	1 n	5 5	R R	5 5
Wt-5	Germany	1 n	1 n	3 1	R R	3 1
Zdr-1	Czech	2 2	2 2	2 2	n R	4 4
Zdr-6	Czech	3 3	3 3	1 1	R R	3 3

CONCLUSIONS AND OUTLOOKS

CONCLUSIONS AND OUTLOOKS

Plant immunity is innate, encoded in the germ line, yet it enables recognition of diverse, often rapidly evolving pathogen-derived molecules. The first two chapters of this Dissertation provide a molecular explanation of this paradox, examining interactions between diverse alleles of an oomycete effector, *ATR1*, and of a cognate *Arabidopsis* resistance gene, *RPP1*. In the first chapter of this dissertation, I show that the recognition of ATR1 by RPP1 is perfectly correlated with ability of the two cognate protein variants to physically interact *in planta*. The interaction was mapped to the Leucine Rich Repeat (LRR) domain of RPP1, which was sufficient for association with ATR1, but not for induction of downstream signaling. Both *ATR1* and *RPP1* genes have been evolving under strong levels of diversifying selection, providing valuable information for the genetic analysis. Using naturally occurring polymorphisms as the basis for mutational analyses, I was able to deduce amino acid sites that were critical for ATR1 recognition. Mutations in those sites converted previously unrecognized alleles to be recognized and vice versa. Interestingly, the two RPP1 variants, Nd1 and WsB, seemed to depend on different, non-overlapping sets of critical residues in ATR1. Obtaining the three dimensional x-ray crystal structure of ATR1 in collaboration with Tom Alber's group was a major advancement that allowed interpreting our mutational data. The amino acid sites in ATR1 that we showed were critical for its recognition by two RPP1 variants that formed distinct surfaces on the protein structure. This allowed drawing a major conclusion: the LRR domains of different RPP1 variants could have evolved separate, independent ways to recognize ATR1. This clearly shows that the LRR domain in plant R-proteins can play a dynamic adaptive role. Being able to use natural polymorphisms in ATR1 to change its recognition specificity suggested that the same strategy could be applied to the LRR domain of RPP1. Indeed, using structural modeling of the LRR, mapping of the natural polymorphisms, and site directed mutagenesis, I was able to identify at least one set of mutations that converted RPP1 to expand its range of recognition. This data confirms our conclusions about the adaptive nature of the LRR and strongly suggests that the nature of ATR1 recognition is through direct binding of two proteins. Most importantly, the ability to put gain-of-recognition mutations in the LRR opens the possibility to genetically engineer plant disease resistance to recognize novel specificities.

Additional data presented in the first chapter allows the understanding of the molecular events downstream of ATR1 recognition. Mutations in the Toll Interleukin 1 (TIR) and Nucleotide Binding Site (NBS) domains of RPP1 abolish the induction of plant defense responses, but do not affect interaction with ATR1. This clearly shows that the activity of those two domains lies downstream of effector recognition. Further information is gained from the ability of the TIR domain alone to induce effector-independent, autoactive defense response. This shows that the activity of the TIR domain is sufficient for induction of a signaling cascade, but is tightly controlled in the context of the full-length protein. The control is probably due to both NBS and the LRR domains, and likely involves intramolecular interactions, nucleotide hydrolysis and dimerization or oligomerization.

Altogether, this information provides a coherent and consistent model, explaining functions of different RPP1 domains, which could be widely applied to other R-proteins.

The first two chapters provide a molecular explanation for gene-for-gene interactions pointing at the adaptive role of the LRR. To gain further understanding of plant immunity, we need to look at its development on the population level. Since plant immunity is encoded in the germ line, the best way to understand the evolution and selection of its components happen is by looking at the genotypes and phenotypes of populations. The third chapter of this dissertation presents phenotypic characterization of *Hpa* / *Arabidopsis* interactions, using 83 Arabidopsis ecotypes and 5 *Hpa* strains, as well as individual alleles of ATR1. Together with a very bright undergraduate student, Connie Zheng, I have been able to show that six *Arabidopsis* ecotypes are able to recognize ATR1. Interestingly, overall genomic divergence that was previously estimated by single nucleotide polymorphism data shows that those ecotypes do not form a single phylogenetic clade. This could be explained in two ways. Either some lineages lost recognition of ATR1, or it was independently acquired. Our structural data from chapter two strengthens the latter hypothesis. Independent generation of effector recognition is also supported by previously published data on another *Hpa* effector, ATR13 [115], where it was shown that recognition of ATR13 in different *Arabidopsis* ecotypes can be achieved by two independent loci, and that the locus recognizing ATR13 can also recognize another unrelated effector. Therefore, it could be predicted that in case of recognition of rapidly evolving effectors, such as ATR1 and ATR13, recognition is the result of an evolutionary teamwork of many R-gene loci. This challenges the existence of truly cognate effector/R-gene interaction. Instead, it points out that this interaction can be viewed as gene-for-gene only at any given snapshot of evolution. In reality, the interactions are defined on genome-by-genome basis, where non-orthologous genes can acquire recognition of a specific effector molecule and recognize it in different ways. This puts forward a problem of nomenclature. If recognition of individual effectors on a population level is not necessary a due to any single locus, but due a physical reality, should any anti-ATR1 R-gene necessarily be called a functional RPP1? And what if the same locus, such as RPP13, is able to recognize two different effectors and vice versa, two different loci recognize the same effector? I can hardly come up with a good answer rather than calling R-genes based on their recognition properties, similar to the antibodies, such as anti-ATR1. This type of "promiscuous recognition" also brings our attention back to the Red Queen Hypothesis, which points at dynamicity of host/pathogen interactions. In case of obligate pathogens, evolutionary pressure is ever stronger since the only pathogen that survives is the one that have been able to successfully escape host immunity, and the host needs to keep up with evolutionary challenge.

In regard to the immune system of vertebrates, the term acquired immunity refers to the generation of vast potential for pathogen recognition that evolves anew in each individual. It is often contrasted with innate immunity that is more evolutionary stable and is based on the recognition of common pathogen associated molecular patterns (PAMPs). In plants, the immunity is usually subdivided into PAMP-triggered immunity and effector-triggered immunity. Collectively, the two branches of plant immunity are still called innate, due to the lack of somatic recombination or development of memory. The data presented in

this dissertation clearly suggests that there is a branch of the effector-triggered immunity in plants that fulfils the role of adaptive immunity in vertebrates. The role of adaptive receptors is fulfilled by rapidly evolving LRR domains of multiple R-genes, which enables recognition of diverse pathogen-derived molecules. There are multiple questions that remained unanswered. First, what is the full recognition potential of any given R-gene: is it possible to find or create an LRR that would recognize a novel protein that is not associated with a pathogen? Second, the exact mechanism of R-protein autoinhibition and its release upon effector-triggered activation remain elusive. Finally, the precise sequence of events starting from R-gene activation to inhibition of pathogen growth remains undefined, although many individual components in this cascade have been determined. Answering these questions will provide an updated view of an ancient immune system and can provide parallels between pathogen defense strategies, developed in different kingdoms of life.

BIBLIOGRAPHY

1. Staskawicz, B.J., et al., *Common and contrasting themes of plant and animal diseases*, in *Science*. 2001. p. 2285-9.
2. Nürnberger, T., et al., *Innate immunity in plants and animals: striking similarities and obvious differences*. Immunol Rev, 2004. **198**: p. 249-66.
3. Van der Biezen, E.A. and J.D. Jones, *Plant disease-resistance proteins and the gene-for-gene concept*. Trends Biochem Sci, 1998. **23**(12): p. 454-6.
4. Chisholm, S.T., et al., *Host-microbe interactions: shaping the evolution of the plant immune response*, in *Cell*. 2006. p. 803-14.
5. Flor, H.H., *Inheritance of pathogenicity in Melampsora lini*. Phytopathology, 1942. **32**: p. 653-69.
6. Flor, H.H., *Current status of the gene-for-gene concept*. Annual Review of Phytopathology, 1971: p. 275-296.
7. Hedrick, S.M., *The acquired immune system: a vantage from beneath*. Immunity, 2004. **21**(5): p. 607-15.
8. Carroll, L., J. Tenniel, and C.L. Dodgson, *Alice's adventures in Wonderland*. 1865, London: Macmillan and co. 6 p.l., 192 p.
9. Belkhadir, Y., R. Subramaniam, and J.L. Dangl, *Plant disease resistance protein signaling: NBS-LRR proteins and their partners*. Curr Opin Plant Biol, 2004. **7**(4): p. 391-9.
10. Coates, M.E. and J.L. Beynon, *Hyaloperonospora Arabidopsidis as a pathogen model*. Annu Rev Phytopathol, 2010. **48**: p. 329-45.
11. Baxter, L., et al., *Signatures of adaptation to obligate biotrophy in the Hyaloperonospora arabidopsidis genome*. Science, 2010. **330**(6010): p. 1549-51.
12. Ausubel, F.M., *Arabidopsis genome. A milestone in plant biology*. Plant Physiology (Rockville), 2000. **124**(4): p. 1451-1454.
13. Holub, E.B., J.L. Beynon, and I.R. Crute, *Phenotypic and genotypic characterization of interactions between isolates of Peronospora parasitica and accessions of Arabidopsis thaliana*. MPMI, 1994. **7**(2): p. 223-239.
14. Bailey, K., et al., *Molecular cloning of ATR5Emoy2 from Hyaloperonospora arabidopsidis, an avirulence determinant that triggers RPP5-mediated defense in Arabidopsis*. Mol Plant Microbe Interact, 2011.
15. Win, J., et al., *Adaptive evolution has targeted the C-terminal domain of the RXLR effectors of plant pathogenic oomycetes*, in *Plant Cell*. 2007. p. 2349-69.
16. Krasileva, K.V., D. Dahlbeck, and B.J. Staskawicz, *Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector*. Plant Cell, 2010. **22**(7): p. 2444-58.
17. Trinchieri, G. and A. Sher, *Cooperation of Toll-like receptor signals in innate immune defence*. Nat Rev Immunol, 2007. **7**(3): p. 179-190.
18. Jones, J.D.G. and J.L. Dangl, *The plant immune system*, in *Nature*. 2006. p. 323-9.

19. Dodds, P.N., et al., *Effectors of biotrophic fungi and oomycetes: pathogenicity factors and triggers of host resistance*. New Phytol, 2009. **183**(4): p. 993-1000.
20. Cornelis, G.R., *The type III secretion injectisome*. Nat Rev Micro, 2006. **4**(11): p. 811-825.
21. Kamoun, S., *A Catalogue of the Effector Secretome of Plant Pathogenic Oomycetes*. Annu. Rev. Phytopathol., 2006. **44**(1): p. 41-60.
22. Desveaux, D., A.U. Singer, and J.L. Dangl, *Type III effector proteins: doppelgangers of bacterial virulence*. Curr Opin Plant Biol, 2006. **9**(4): p. 376-82.
23. Takken, F.L., M. Albrecht, and W.I. Tameling, *Resistance proteins: molecular switches of plant defence*, in *Curr Opin Plant Biol*. 2006. p. 383-90.
24. Kim, H.-E., et al., *Formation of apoptosome is initiated by cytochrome c-induced dATP hydrolysis and subsequent nucleotide exchange on Apaf-1*. Proc Natl Acad Sci USA, 2005. **102**(49): p. 17545-50.
25. Tameling, W.I., et al., *The tomato R gene products I-2 and MI-1 are functional ATP binding proteins with ATPase activity*. Plant Cell, 2002. **14**(11): p. 2929-39.
26. Wirthmueller, L., et al., *Nuclear accumulation of the Arabidopsis immune receptor RPS4 is necessary for triggering EDS1-dependent defense*, in *Curr Biol*. 2007. p. 2023-9.
27. Mestre, P. and D.C. Baulcombe, *Elicitor-mediated oligomerization of the tobacco N disease resistance protein*, in *Plant Cell*. 2006. p. 491-501.
28. Swiderski, M.R., D. Birker, and J.D.G. Jones, *The TIR Domain of TIR-NB-LRR Resistance Proteins Is a Signaling Domain Involved in Cell Death Induction*. Molecular Plant-Microbe Interactions, 2009. **22**(2): p. 157-165.
29. Burch-Smith, T.M., et al., *A novel role for the TIR domain in association with pathogen-derived elicitors*, in *Plos Biol*. 2007. p. e68.
30. Ellis, J.G., et al., *Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity*. Plant Cell, 1999. **11**(3): p. 495-506.
31. Dinesh-Kumar, S.P., W.H. Tham, and B.J. Baker, *Structure-function analysis of the tobacco mosaic virus resistance gene N*. Proc Natl Acad Sci U S A, 2000. **97**(26): p. 14789-94.
32. Kobe, B. and J. Deisenhofer, *Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats*. Nature, 1993. **366**(6457): p. 751-6.
33. Alder, M.N., et al., *Diversity and function of adaptive immune receptors in a jawless vertebrate*. Science, 2005. **310**(5756): p. 1970-3.
34. Catanzariti, A.-M., et al., *The AvrM effector from flax rust has a structured C-terminal domain and interacts directly with the M resistance protein*, in *Molecular Plant-Microbe Interactions*. 2010. p. 49-57.
35. Dodds, P.N., et al., *Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes*. Proc Natl Acad Sci U S A, 2006. **103**(23): p. 8888-93.
36. Deslandes, L., et al., *Physical interaction between RRS1-R, a protein conferring resistance to bacterial wilt, and PopP2, a type III effector targeted to the plant nucleus*. Proc Natl Acad Sci U S A, 2003. **100**(13): p. 8024-9.

37. Mondragon-Palomino, M., et al., *Patterns of positive selection in the complete NBS-LRR gene family of Arabidopsis thaliana*. Genome Res, 2002. **12**(9): p. 1305-15.
38. Allen, R.L., et al., *Host-parasite coevolutionary conflict between Arabidopsis and downy mildew*, in *Science*. 2004. p. 1957-60.
39. Rentel, M.C., et al., *Recognition of the Hyaloperonospora parasitica effector ATR13 triggers resistance against oomycete, bacterial, and viral pathogens*. Proc Natl Acad Sci USA, 2008. **105**(3): p. 1091-6.
40. Shen, Q.-H., et al., *Nuclear activity of MLA immune receptors links isolate-specific and basal disease-resistance responses*, in *Science*. 2007. p. 1098-103.
41. Rairdan, G.J. and P. Moffett, *Distinct domains in the ARC region of the potato resistance protein Rx mediate LRR binding and inhibition of activation*. Plant Cell, 2006. **18**(8): p. 2082-93.
42. Dodds, P.N., G.J. Lawrence, and J.G. Ellis, *Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference between the P and P2 rust resistance specificities in flax*. Plant Cell, 2001. **13**(1): p. 163-78.
43. Jia, Y., et al., *Direct interaction of resistance gene and avirulence gene products confers rice blast resistance*. EMBO (European Molecular Biology Organization) Journal, 2000. **19**(15): p. 4004-4014.
44. Botella, M.A., et al., *Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct Peronospora parasitica avirulence determinants*. The Plant cell, 1998. **10**: p. 1847-1860.
45. Rehmany, A.P., et al., *Differential recognition of highly divergent downy mildew avirulence gene alleles by RPP1 resistance genes from two Arabidopsis lines*, in *Plant Cell*. 2005. p. 1839-50.
46. Sohn, K.H., et al., *The downy mildew effector proteins ATR1 and ATR13 promote disease susceptibility in Arabidopsis thaliana*. Plant Cell, 2007. **19**(12): p. 4077-90.
47. Bomblies, K., et al., *Autoimmune response as a mechanism for a Dobzhansky-Muller-type incompatibility syndrome in plants*, in *Plos Biol*. 2007. p. e236.
48. Birch, P.R., et al., *Trafficking arms: oomycete effectors enter host plant cells*. Trends Microbiol, 2006. **14**(1): p. 8-11.
49. Zhang, Y., et al., *Expression of RPS4 in tobacco induces an AvrRps4-independent HR that requires EDS1, SGT1 and HSP90*. Plant J, 2004. **40**(2): p. 213-24.
50. Weaver, M.L., et al., *The Arabidopsis thaliana TIR-NB-LRR R-protein, RPP1A; protein localization and constitutive activation of defence by truncated alleles in tobacco and Arabidopsis*, in *Plant J*. 2006. p. 829-40.
51. Tameling, W.I.L., *Mutations in the NB-ARC Domain of I-2 That Impair ATP Hydrolysis Cause Autoactivation*. PLANT PHYSIOLOGY, 2006. **140**(4): p. 1233-1245.
52. Frost, D., et al., *Tobacco transgenic for the flax rust resistance gene L expresses allele-specific activation of defense responses*. Mol Plant Microbe Interact, 2004. **17**(2): p. 224-32.

53. Shaner, N.C., P.A. Steinbach, and R.Y. Tsien, *A guide to choosing fluorescent proteins*. Nat Methods, 2005. **2**(12): p. 905-9.
54. Horton, P., et al., *WoLF PSORT: protein localization predictor*. Nucleic Acids Res, 2007. **35**(Web Server issue): p. W585-7.
55. Mucyn, T.S., et al., *The tomato NBARC-LRR protein Prf interacts with Pto kinase in vivo to regulate specific plant immunity*, in *Plant Cell*. 2006. p. 2792-806.
56. Axtell, M.J. and B.J. Staskawicz, *Initiation of RPS2-specified disease resistance in Arabidopsis is coupled to the AvrRpt2-directed elimination of RIN4*, in *Cell*. 2003. p. 369-77.
57. Shao, F., et al., *Cleavage of Arabidopsis PBS1 by a bacterial type III effector*. Science, 2003. **301**(5637): p. 1230-3.
58. Ade, J., et al., *Indirect activation of a plant nucleotide binding site-leucine-rich repeat protein by a bacterial protease*. Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2531-6.
59. Mackey, D., et al., *RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis*, in *Cell*. 2002. p. 743-54.
60. Dangl, J.L. and J.M. McDowell, *Two modes of pathogen recognition by plants*. Proc Natl Acad Sci U S A, 2006. **103**(23): p. 8575-6.
61. Barrett, L.G., et al., *Diversity and Evolution of Effector Loci in Natural Populations of the Plant Pathogen Melampsora lini*, in *Mol Biol Evol*. 2009. p. 2499-513.
62. Van der Hoorn, R.A., P.J. De Wit, and M.H. Joosten, *Balancing selection favors guarding resistance proteins*. Trends Plant Sci, 2002. **7**(2): p. 67-71.
63. Rohmer, L., D.S. Guttman, and J.L. Dangl, *Diverse evolutionary mechanisms shape the type III effector virulence factor repertoire in the plant pathogen Pseudomonas syringae*. Genetics, 2004. **167**(3): p. 1341-60.
64. Rairdan, G.J., et al., *The coiled-coil and nucleotide binding domains of the Potato Rx disease resistance protein function in pathogen recognition and signaling*. Plant Cell, 2008. **20**(3): p. 739-51.
65. Ueda, H., Y. Yamaguchi, and H. Sano, *Direct interaction between the tobacco mosaic virus helicase domain and the ATP-bound resistance protein, N factor during the hypersensitive response in tobacco plants*, in *Plant Mol Biol*. 2006. p. 31-45.
66. Collier, S.M. and P. Moffett, *NB-LRRs work a "bait and switch" on pathogens*, in *Trends Plant Sci*. 2009. p. 521-9.
67. Burch-Smith, T.M. and S.P. Dinesh-Kumar, *The Functions of Plant TIR Domains*. Science's STKE, 2007. **2007**(401): p. pe46-pe46.
68. Lukasik, E. and F.L.W. Takken, *STANDing strong, resistance proteins instigators of plant defence*. Current Opinion in Plant Biology, 2009. **12**(4): p. 427-36.
69. Rairdan, G.J., *Distinct Domains in the ARC Region of the Potato Resistance Protein Rx Mediate LRR Binding and Inhibition of Activation*. Plant Cell, 2006. **18**(8): p. 2082-2093.

70. Moffett, P., et al., *Interaction between domains of a plant NBS-LRR protein in disease resistance-related cell death*. EMBO J., 2002. **21**(17): p. 4511-9.
71. Edgar, R., *MUSCLE: multiple sequence alignment with high accuracy and high throughput*. Nucleic Acids Research, 2004.
72. Sonnhammer, E.L. and V. Hollich, *Scoredist: a simple and robust protein sequence distance estimator*. BMC Bioinformatics, 2005. **6**: p. 108.
73. Suyama, M., D. Torrents, and P. Bork, *PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments*. Nucleic Acids Research, 2006. **34**(Web Server issue): p. W609-12.
74. Felsenstein, J., *Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods*. Methods Enzymol, 1996. **266**: p. 418-27.
75. Page, R.D., *TreeView: an application to display phylogenetic trees on personal computers*. Comput Appl Biosci, 1996. **12**(4): p. 357-8.
76. Wise, A.A., Z. Liu, and A.N. Binns, *Three methods for the introduction of foreign DNA into Agrobacterium*. Methods Mol Biol, 2006. **343**: p. 43-53.
77. Earley, K.W., et al., *Gateway-compatible vectors for plant functional genomics and proteomics*. The Plant Journal, 2006. **45**(4): p. 616-629.
78. Göker, M., et al., *How do obligate parasites evolve? A multi-gene phylogenetic analysis of downy mildews*, in *Fungal Genet Biol*. 2007. p. 105-22.
79. Bhattacharjee, S., et al., *The Malarial Host-Targeting Signal Is Conserved in the Irish Potato Famine Pathogen*. PLoS Pathog, 2006. **2**(5): p. e50.
80. Dou, D., et al., *Conserved C-terminal motifs required for avirulence and suppression of cell death by Phytophthora sojae effector Avr1b*. Plant Cell, 2008. **20**(4): p. 1118-33.
81. Grouffaud, S., et al., *Plasmodium falciparum and Hyaloperonospora parasitica effector translocation motifs are functional in Phytophthora infestans*. Microbiology, 2008. **154**(12): p. 3743-3751.
82. Whisson, S.C., et al., *A translocation signal for delivery of oomycete effector proteins into host plant cells*. Nature, 2007. **450**(7166): p. 115-118.
83. Jiang, R.H.Y., et al., *RXLR effector reservoir in two Phytophthora species is dominated by a single rapidly evolving superfamily with more than 700 members*, in *Proc Natl Acad Sci USA*. 2008. p. 4874-9.
84. Andersson, M.X., et al., *Phospholipase-dependent signalling during the AvrRpm1- and AvrRpt2-induced disease resistance responses in Arabidopsis thaliana*, in *Plant J*. 2006. p. 947-59.
85. Kobe, B. and J. Deisenhofer, *Proteins with leucine-rich repeats*. Current Opinion in Structural Biology, 1995. **5**: p. 408-416.
86. Wang, C.-I.A., et al., *Crystal structures of flax rust avirulence proteins AvrL567-A and -D reveal details of the structural basis for flax disease resistance specificity*. Plant Cell, 2007. **19**(9): p. 2898-2912.
87. Dong, J., et al., *Crystal structure of the complex between Pseudomonas effector AvrPtoB and the tomato Pto kinase reveals both a shared and a unique interface compared with AvrPto-Pto*. Plant Cell, 2009. **21**(6): p. 1846-59.
88. Lee, C.C., et al., *Crystal structure of the type III effector AvrB from Pseudomonas syringae*. Structure, 2004. **12**(3): p. 487-94.

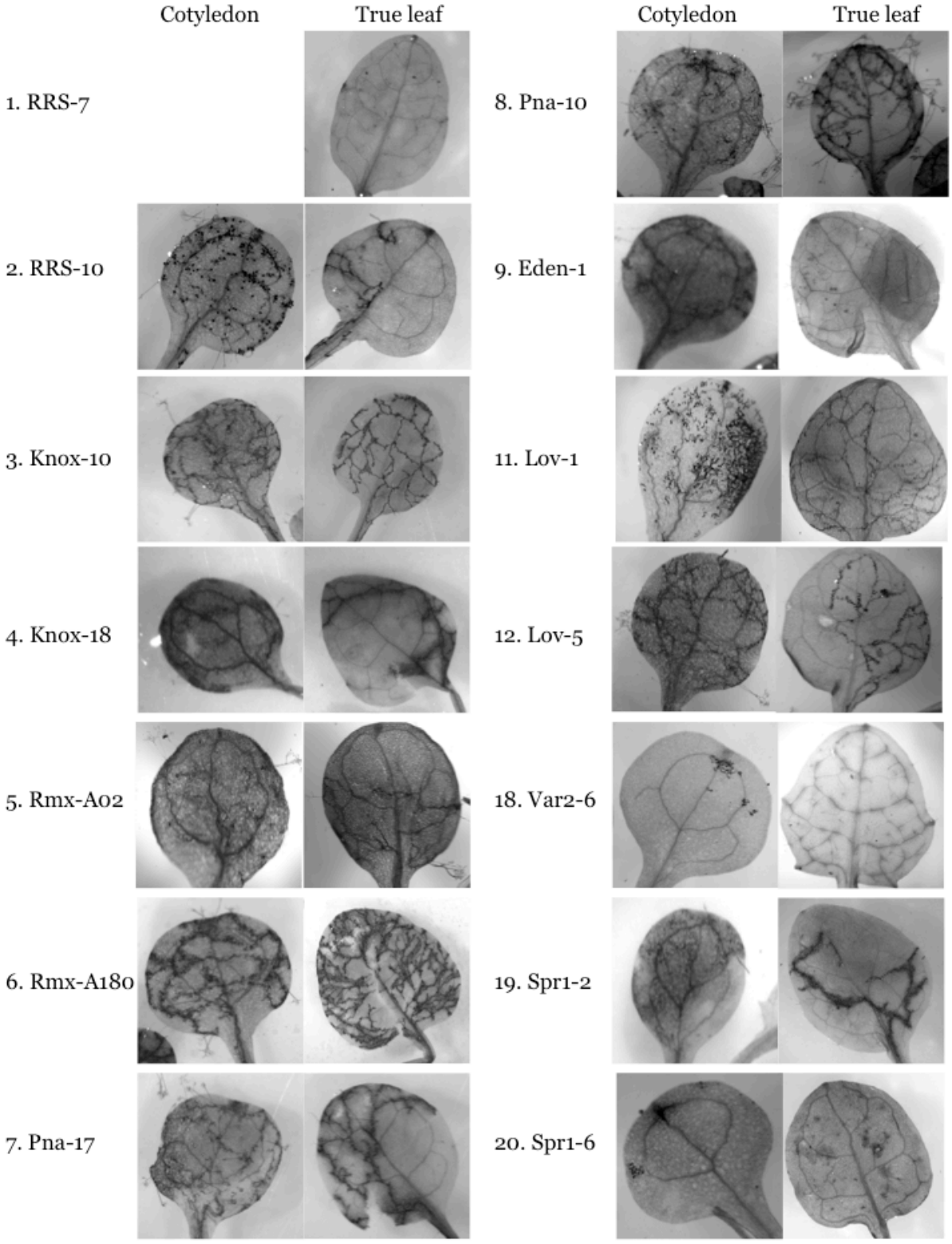
89. Bernoux, M., et al., *Structural and Functional Analysis of a Plant Resistance Protein TIR Domain Reveals Interfaces for Self-Association, Signaling, and Autoregulation*. Cell Host Microbe, 2011. **9**(3): p. 200-11.
90. Williams, S.B., et al., *Structure and function from the circadian clock protein KaiA of Synechococcus elongatus: a potential clock input mechanism*, in *Proc Natl Acad Sci USA*. 2002. p. 15357-62.
91. Wang, W., et al., *Timing of plant immune responses by a central circadian regulator*. Nature. **470**(7332): p. 110-4.
92. Bos, J.I.B., et al., *Distinct amino acids of the Phytophthora infestans effector AVR3a condition activation of R3a hypersensitivity and suppression of cell death*. Mol Plant Microbe Interact, 2009. **22**(3): p. 269-81.
93. MacDowell, A.A., et al., *Suite of three protein crystallography beamlines with single superconducting bend magnet as the source*. J Synchrotron Radiat, 2004. **11**(Pt 6): p. 447-55.
94. Holton, J. and T. Alber, *Automated protein crystal structure determination using ELVES*. Proc Natl Acad Sci USA, 2004. **101**(6): p. 1537-42.
95. Adams, P.D., et al., *PHENIX: building new software for automated crystallographic structure determination*. Acta Crystallogr D Biol Crystallogr, 2002. **58**(Pt 11): p. 1948-54.
96. Zwart, P.H., et al., *Automated structure solution with the PHENIX suite*. Methods Mol Biol, 2008. **426**: p. 419-35.
97. Davis, I.W., et al., *MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes*. Nucleic Acids Res, 2004. **32**(Web Server issue): p. W615-9.
98. Shindyalov, I.N. and P.E. Bourne, *Protein structure alignment by incremental combinatorial extension (CE) of the optimal path*. Protein Eng, 1998. **11**(9): p. 739-47.
99. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis*. J Comput Chem, 2004. **25**(13): p. 1605-12.
100. Holm, L. and P. Rosenstrom, *Dali server: conservation mapping in 3D*. Nucleic Acids Res. **38**(Web Server issue): p. W545-9.
101. Bendtsen, J.D., et al., *Improved prediction of signal peptides: SignalP 3.0*. J Mol Biol, 2004. **340**(4): p. 783-95.
102. Crooks, G.E., et al., *WebLogo: a sequence logo generator*. Genome Res, 2004. **14**(6): p. 1188-90.
103. Koch, E. and A. Slusarenko, *Arabidopsis is susceptible to infection by a downy mildew fungus*, in *Plant Cell*. 1990. p. 437-45.
104. Holub, E., Beynon J., and Crute I., *Phenotypic and genotypic characterization of interactions between isolates of Peronospora parasitica and accessions of Arabidopsis thaliana* Molecular Plant-Microbe Interactions, 1994.
105. Bittner-Eddy, P., et al., *Genetic and physical mapping of the RPP13 locus, in Arabidopsis, responsible for specific recognition of several Peronospora parasitica (downy mildew) isolates*, in *Molecular Plant-Microbe Interactions*. 1999. p. 792-802.

106. Tör, M., et al., *Arabidopsis downy mildew resistance gene RPP27 encodes a receptor-like protein similar to CLAVATA2 and tomato Cf-9*. PLANT PHYSIOLOGY, 2007. **143**(2): p. 1079.
107. Sinapidou, E., et al., *Two TIR:NB:LRR genes are required to specify resistance to Peronospora parasitica isolate Cala2 in Arabidopsis*. Plant J, 2004. **38**(6): p. 898-909.
108. Parker, J.E., et al., *The Arabidopsis downy mildew resistance gene RPP5 shares similarity to the toll and interleukin-1 receptors with N and L6*. The Plant Cell, 1997. **9**: p. 879-894.
109. van der Biezen, E.A., et al., *Arabidopsis RPP4 is a member of the RPP5 multigene family of TIR-NB-LRR genes and confers downy mildew resistance through multiple signalling components*, in *Plant J*. 2002. p. 439-51.
110. Atwell, S., et al., *Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines*. Nature, 2010. **465**(7298): p. 627-31.
111. Holub, E., *Natural history of Arabidopsis thaliana and oomycete symbioses*. European Journal of Plant Pathology, 2008.
112. Nemri, A., et al., *Genome-wide survey of Arabidopsis natural variation in downy mildew resistance using combined association and linkage mapping*. Proc Natl Acad Sci USA, 2010. **107**(22): p. 10302-7.
113. McDowell, J.M., et al., *Genetic analysis of developmentally regulated resistance to downy mildew (Hyaloperonospora parasitica) in Arabidopsis thaliana*, in *Molecular Plant-Microbe Interactions*. 2005. p. 1226-34.
114. Aranzana, M.J., et al., *Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes*. PLoS Genet, 2005. **1**(5): p. e60.
115. Hall, S.A., et al., *Maintenance of genetic variation in plants and pathogens involves complex networks of gene-for-gene interactions*, in *Mol Plant Pathol*. 2009. p. 449-57.
116. Nordborg, M., et al., *The pattern of polymorphism in Arabidopsis thaliana*. Plos Biol, 2005. **3**(7): p. e196.
117. Gou, J.Y., et al., *Negative Regulation of Anthocyanin Biosynthesis in Arabidopsis by a miR156-Targeted SPL Transcription Factor*. Plant Cell.
118. Katagiri, F., R. Thilmony, and S.Y. He, *The Arabidopsis Thaliana-Pseudomonas Syringae Interaction*. The Arabidopsis Book, 2002. **20**(1): p. 1.

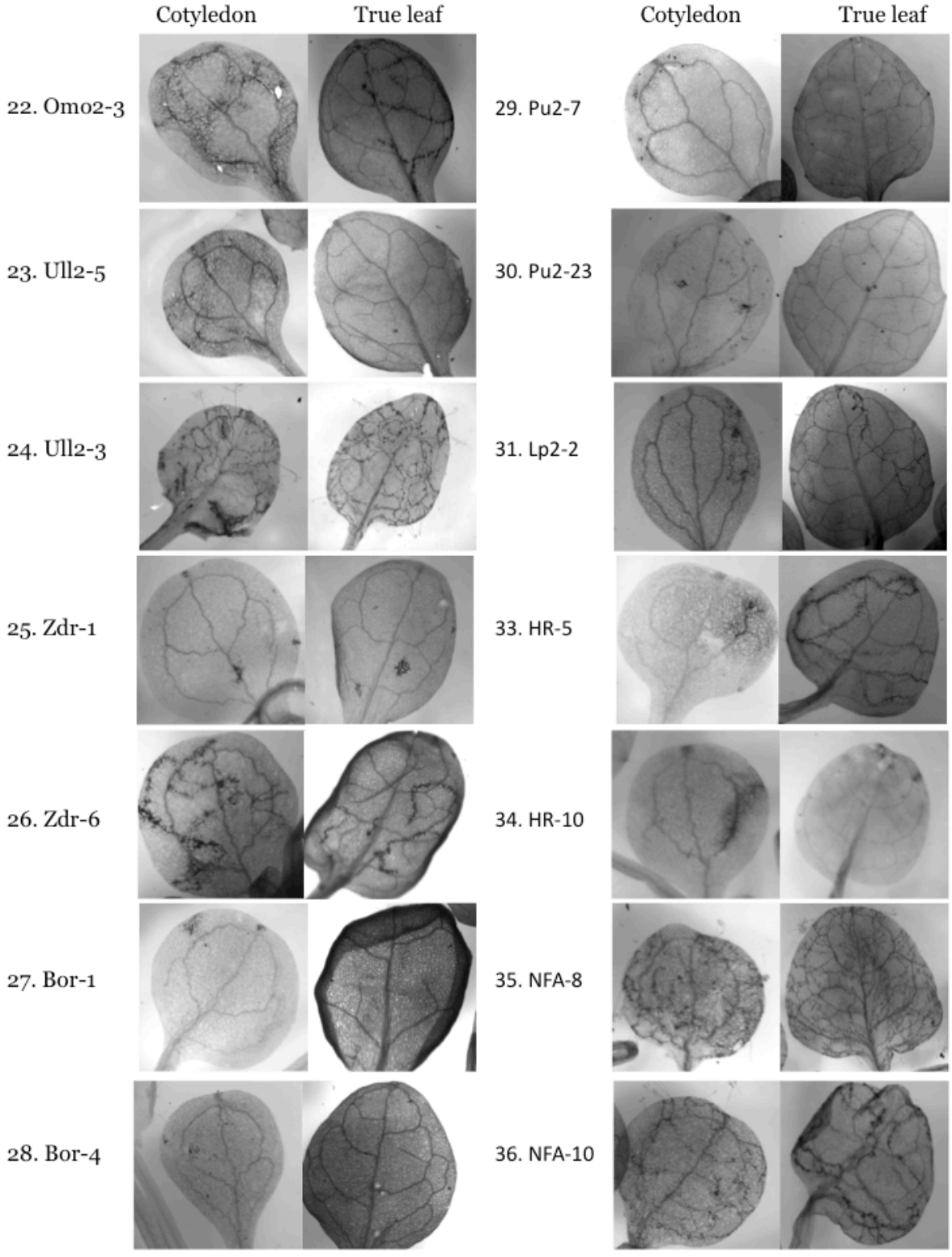
APPENDIX.

Pictures of the *Arabidopsis* cotyledons and true leaves from 83 ecotypes inoculated with *Hpa* strains Emoy2, Emco5 and Emwa1.

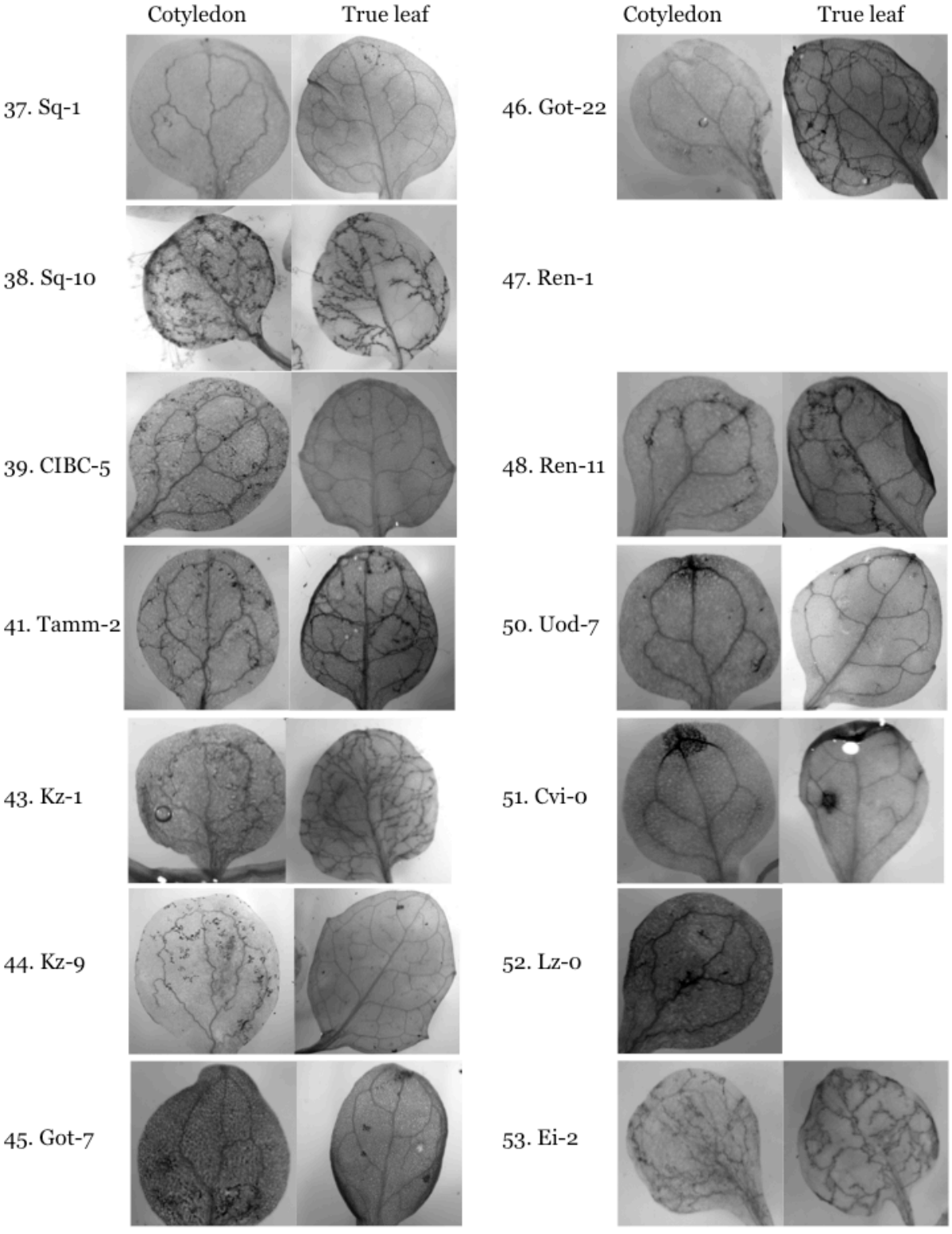
Appendix. Hpa Emoy2



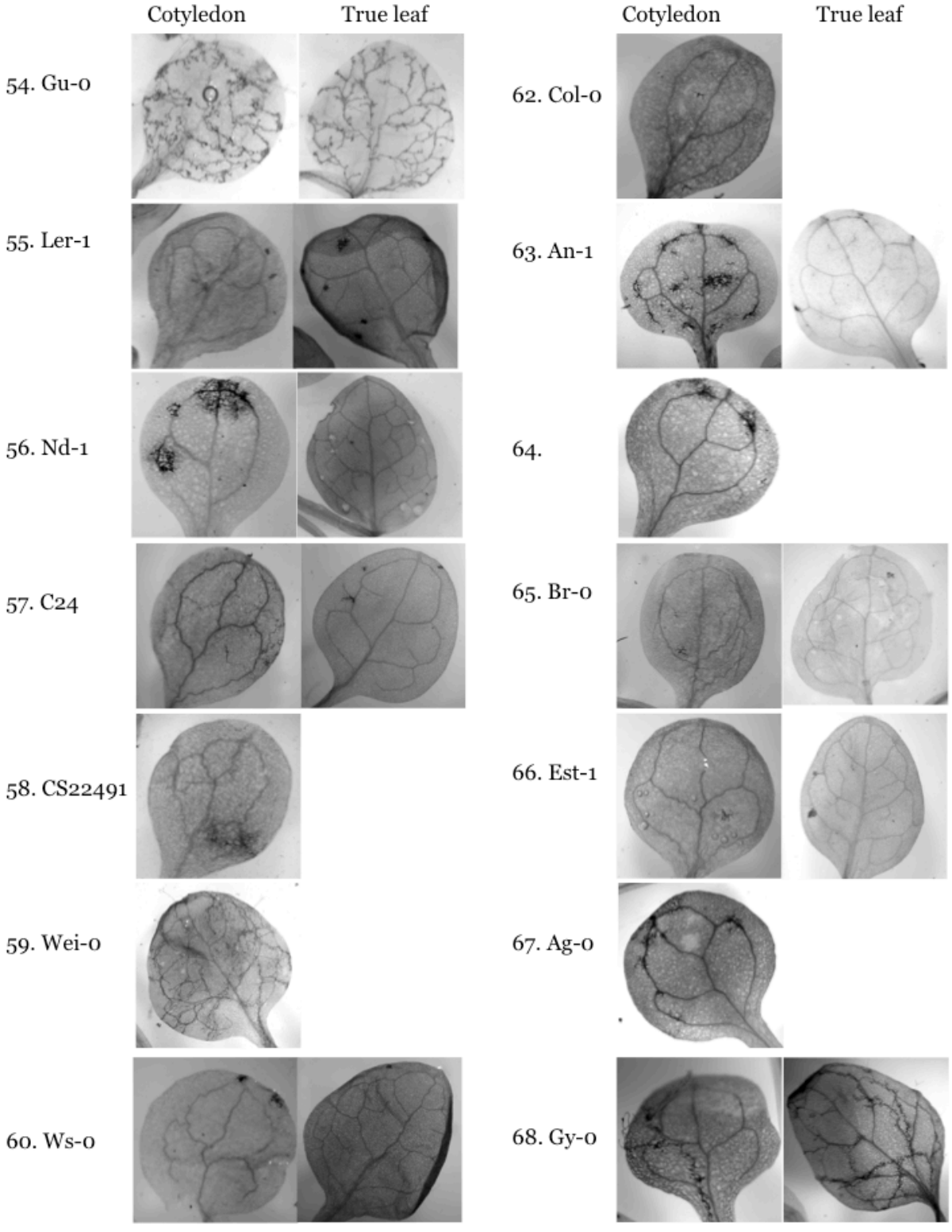
Appendix. Hpa Emoy2 (cont'd)



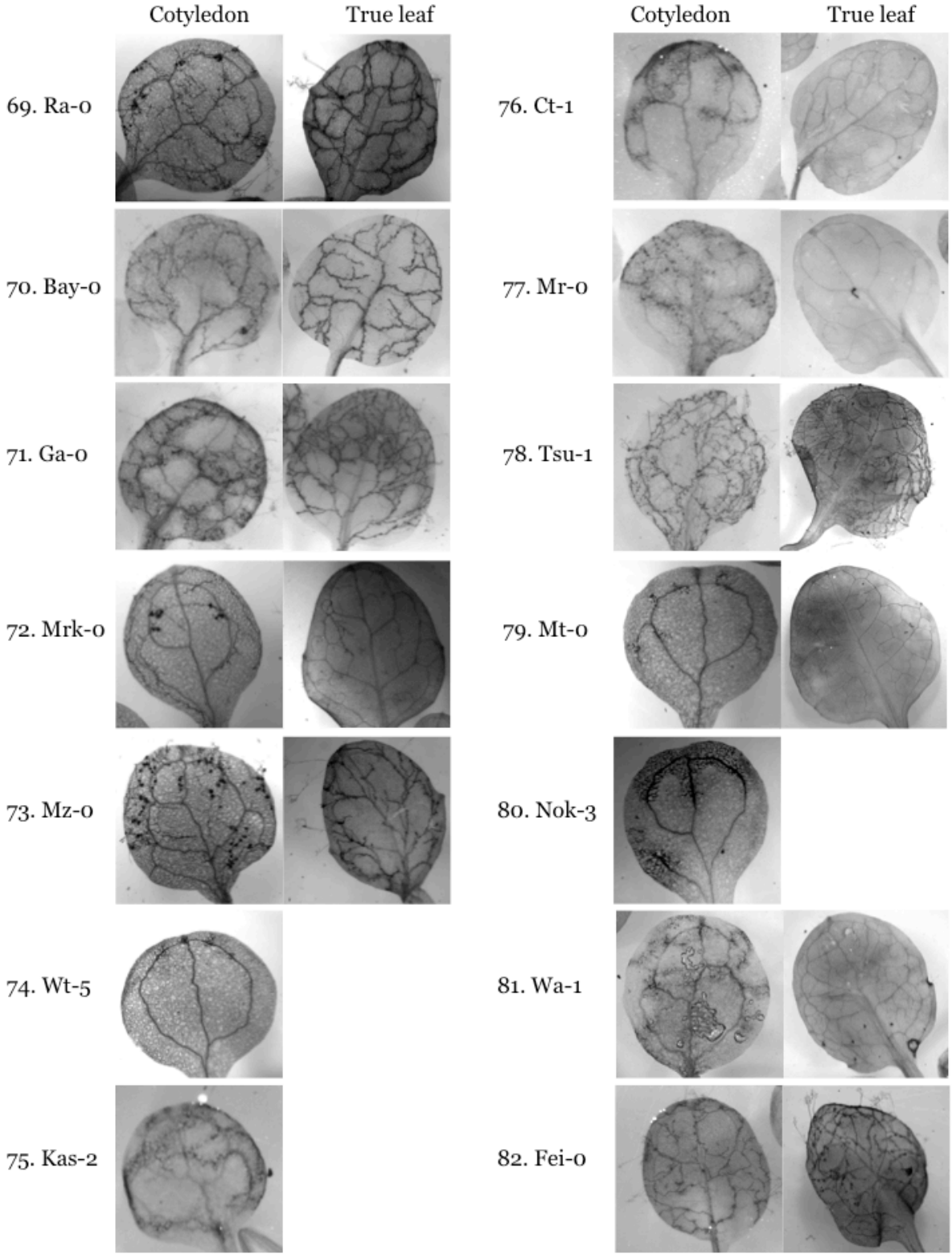
Appendix. Hpa Emoy2 (cont'd)



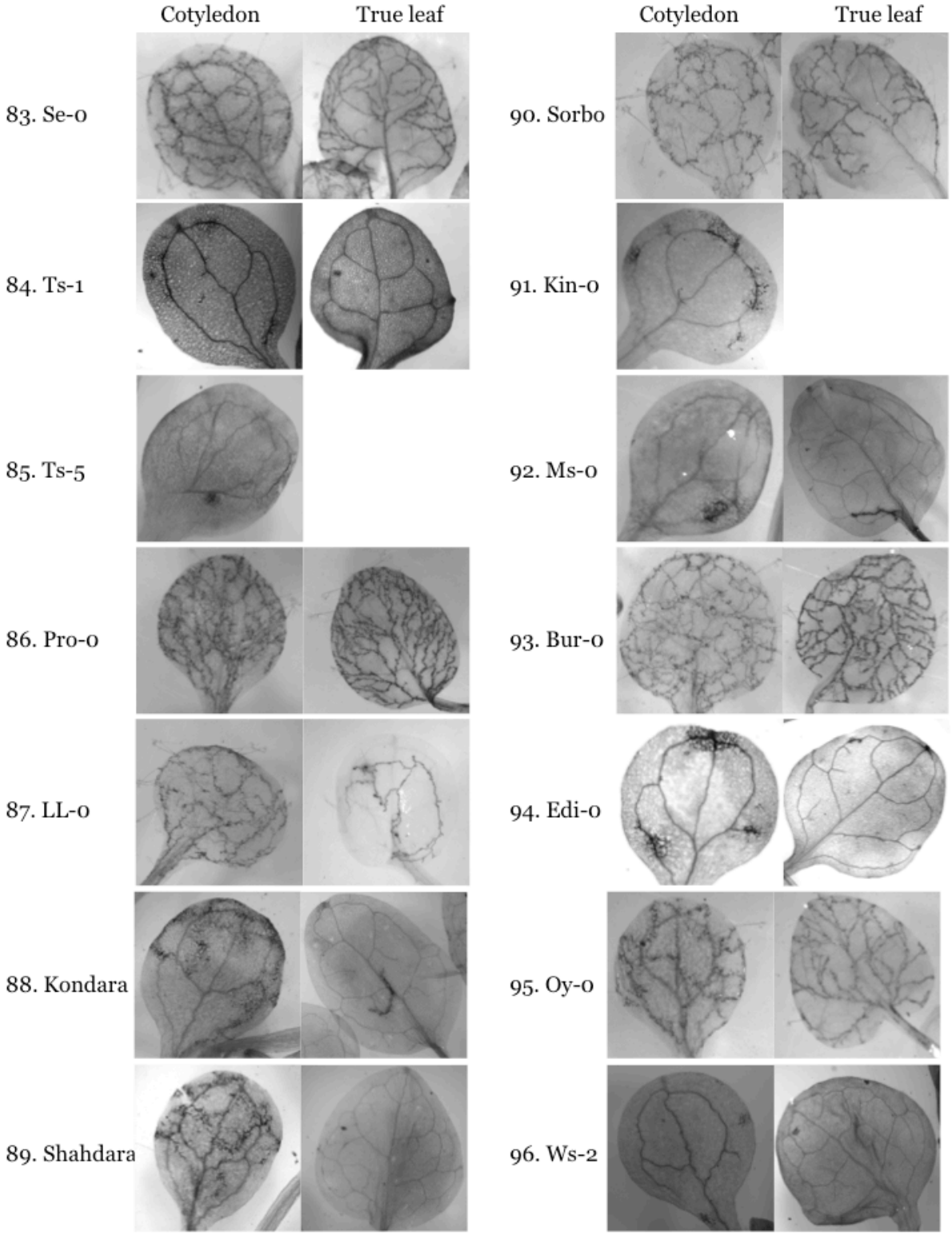
Appendix. Hpa Emoy2 (cont'd)



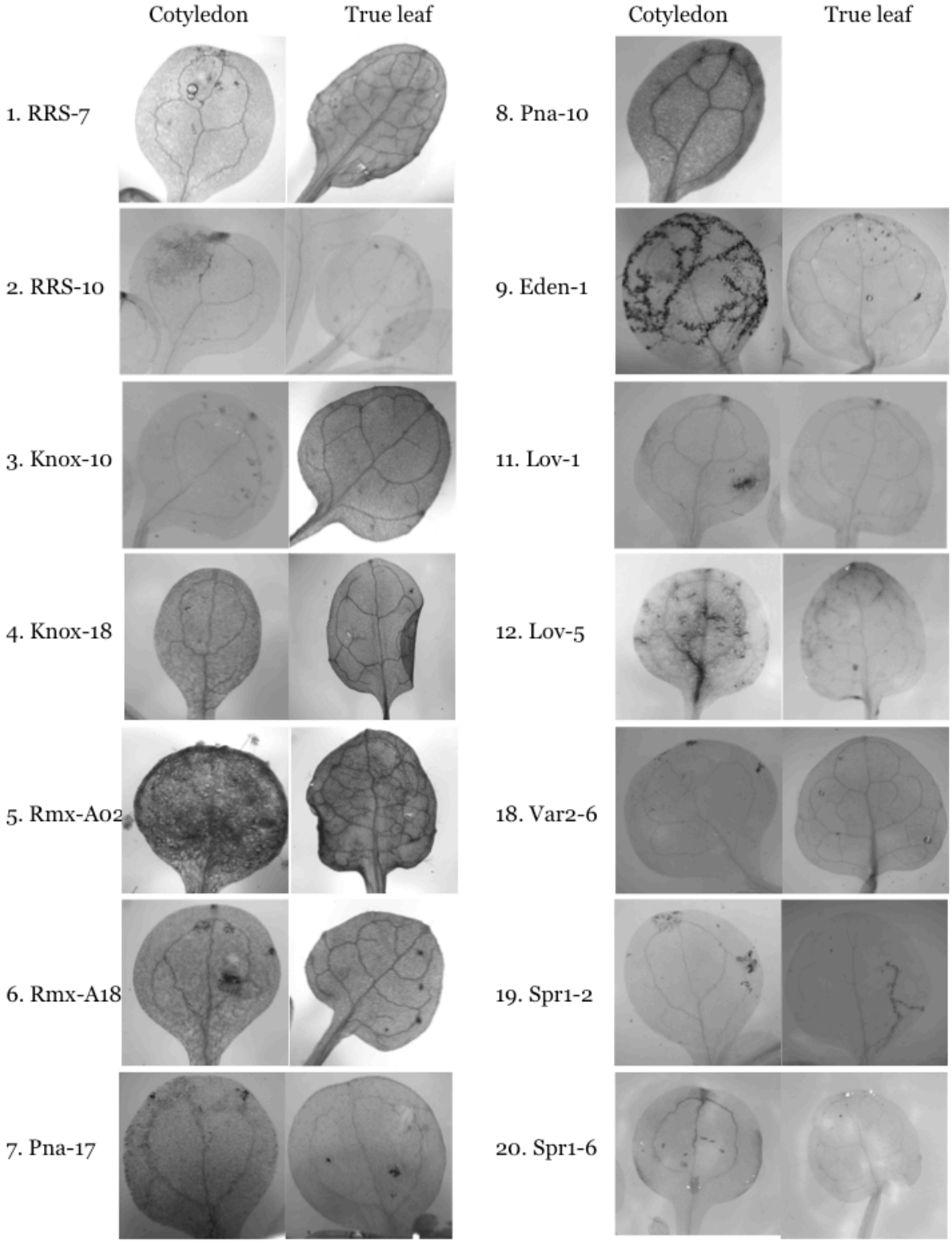
Appendix. Hpa Emoy2 (cont'd)



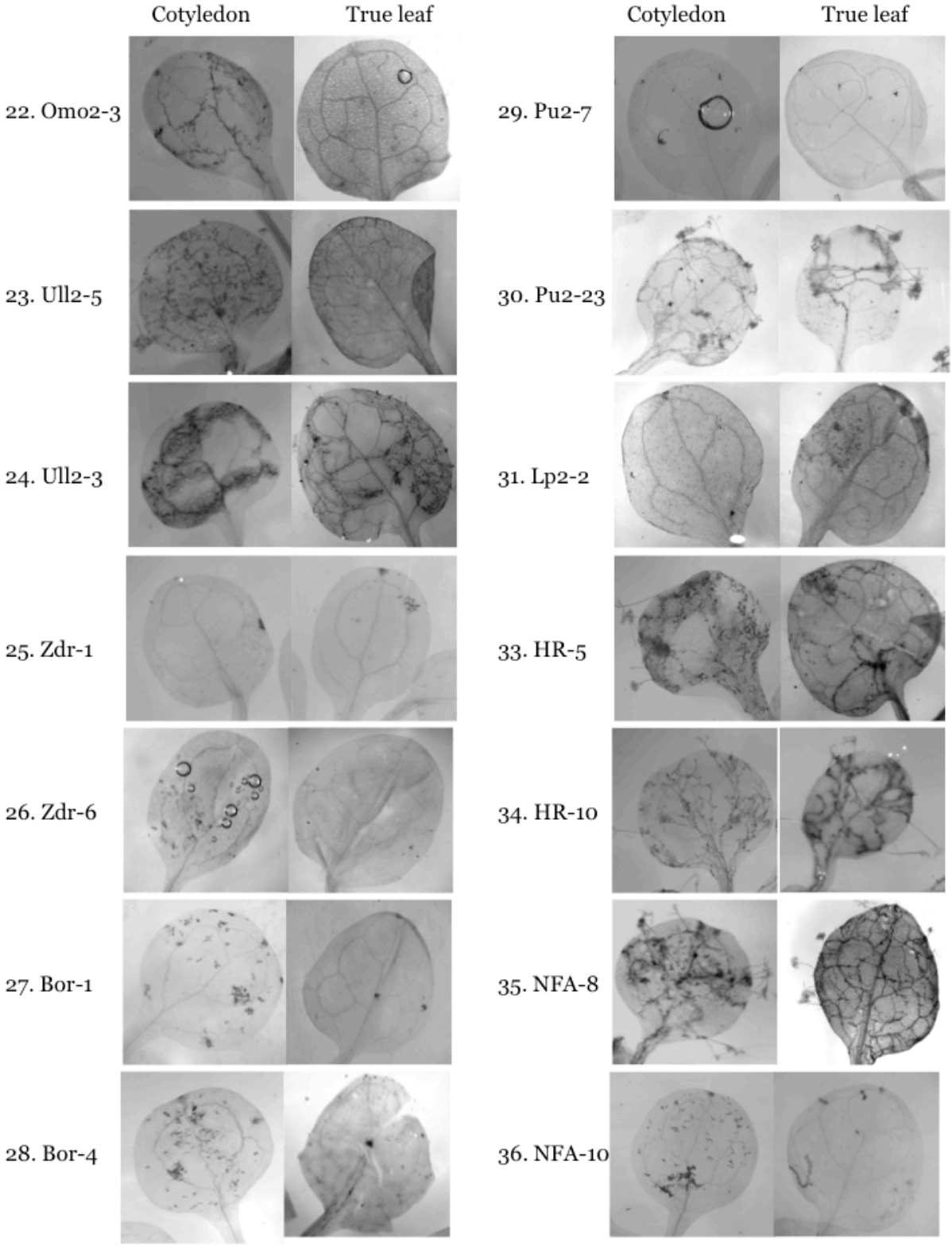
Appendix. Hpa Emoy2 (cont'd)



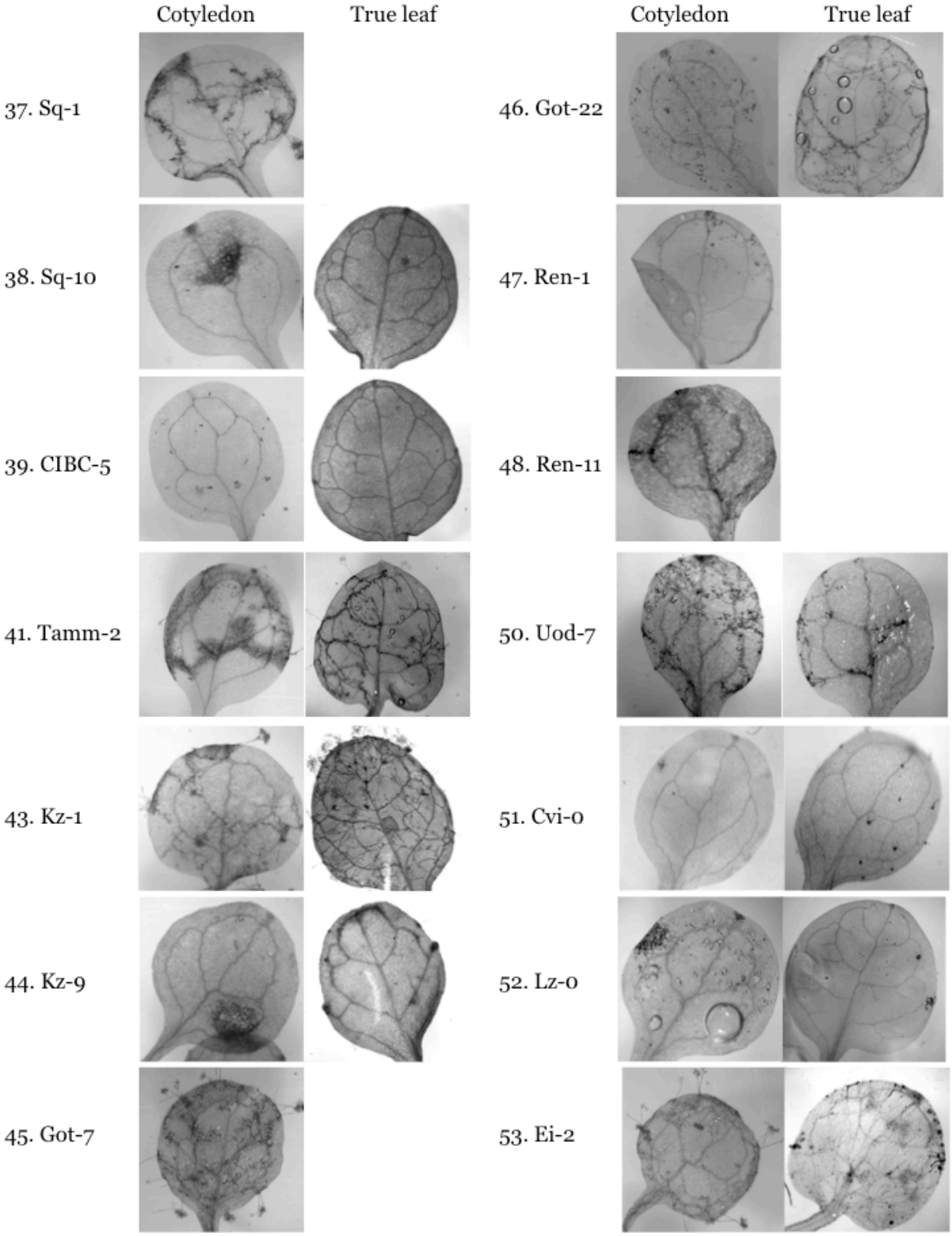
Appendix. Hpa Emco5



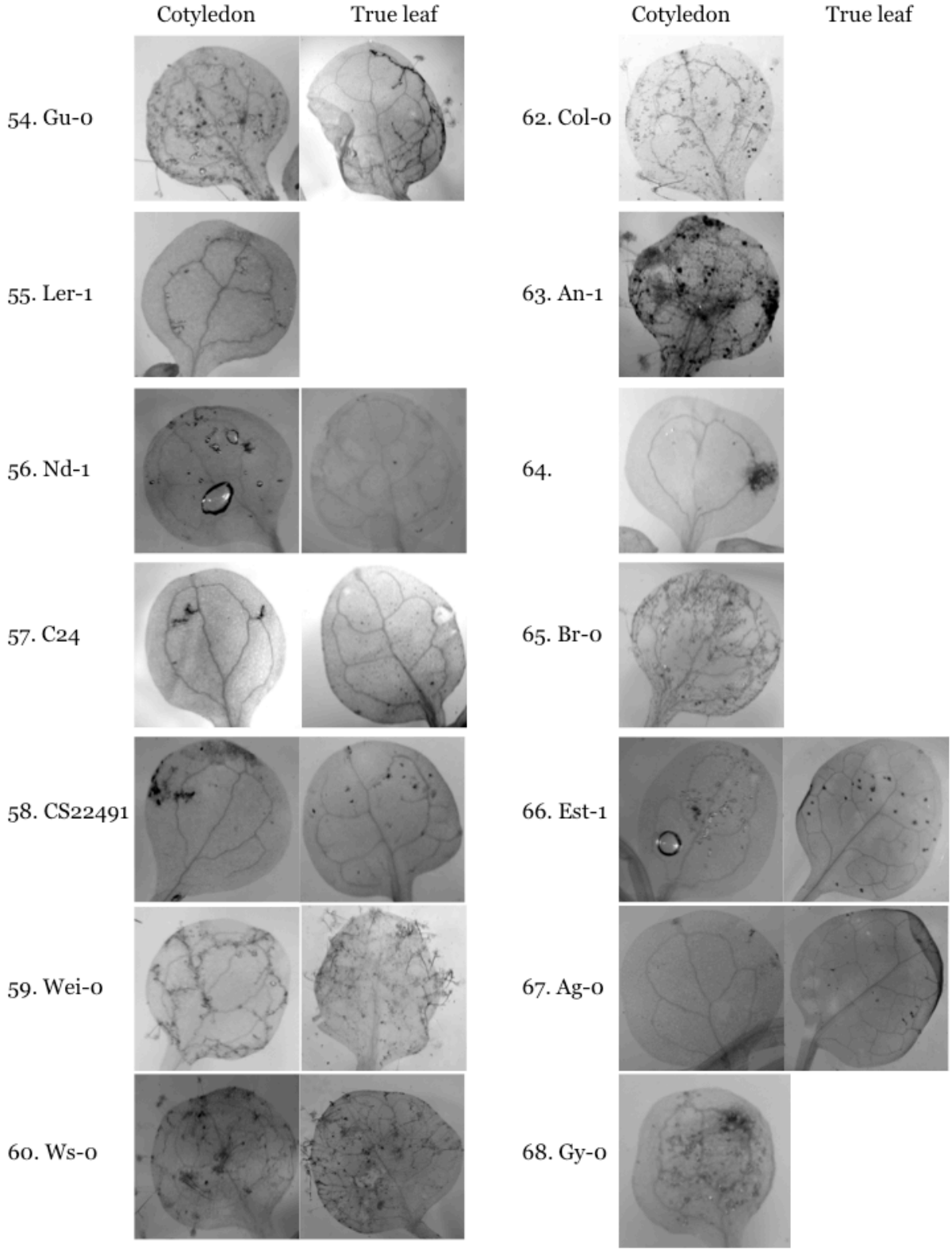
Appendix. Hpa Emco5 (cont'd)



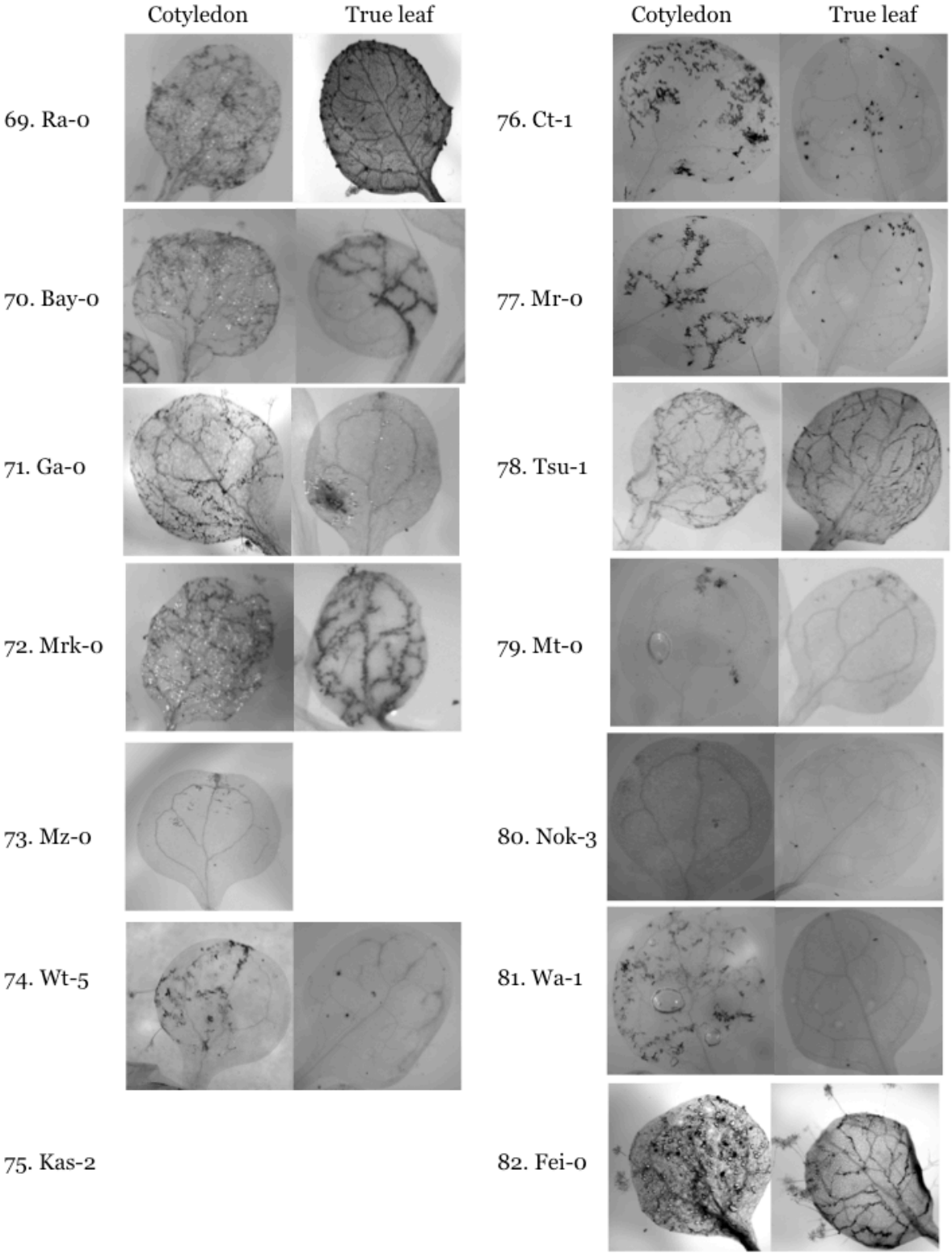
Appendix. Hpa Emco5 (cont'd)



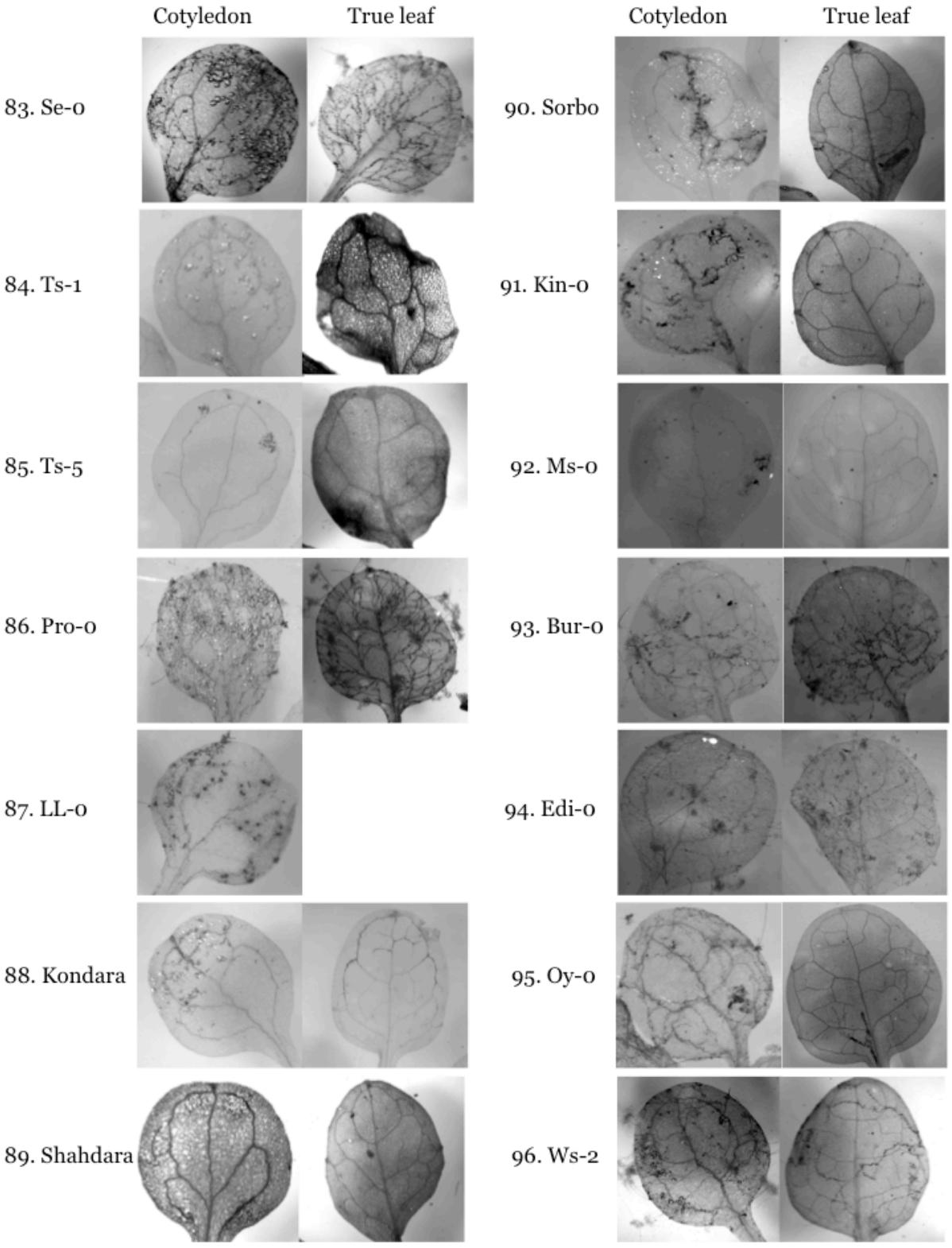
Appendix. Hpa Emco5 (cont'd)



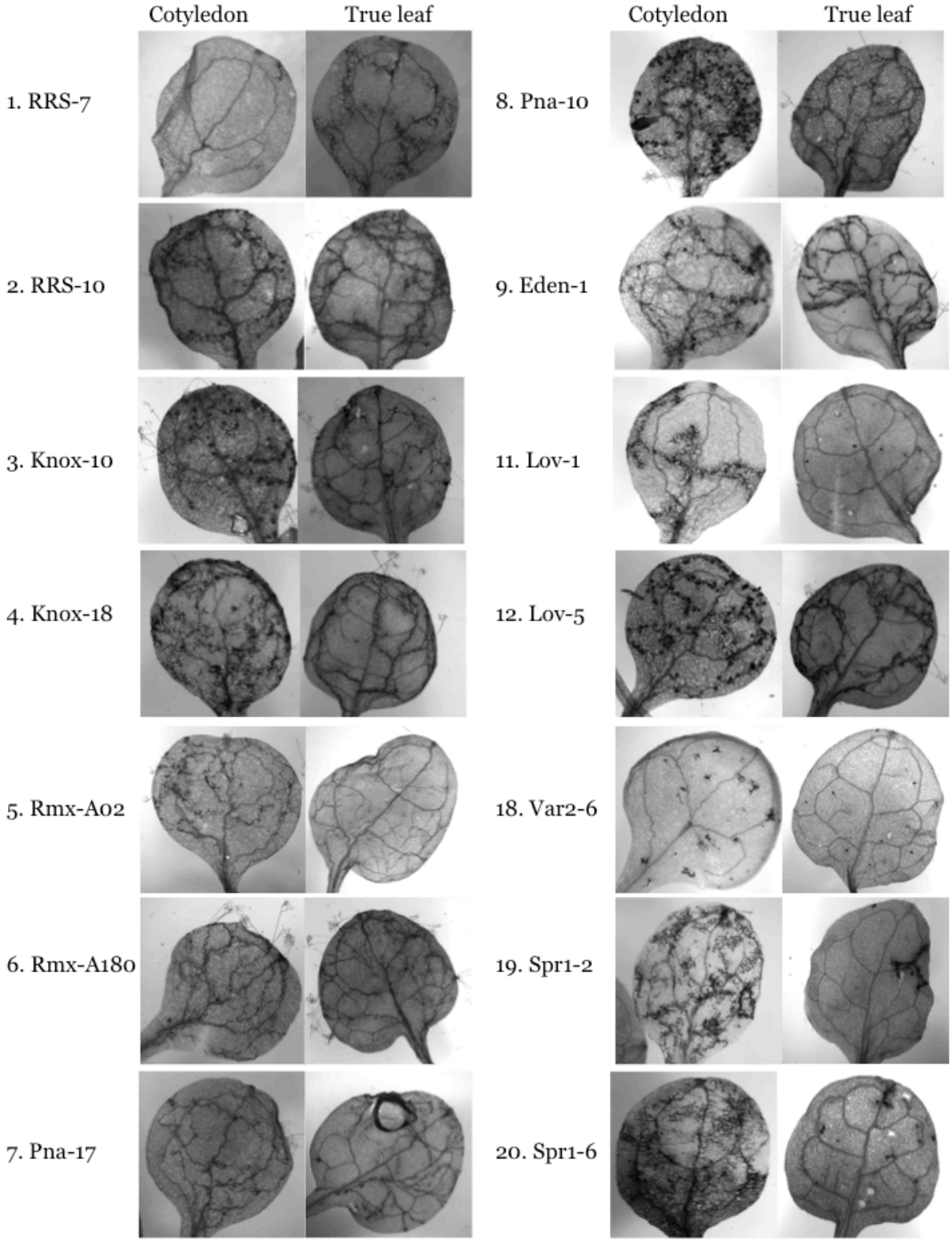
Appendix. Hpa Emco5 (cont'd)



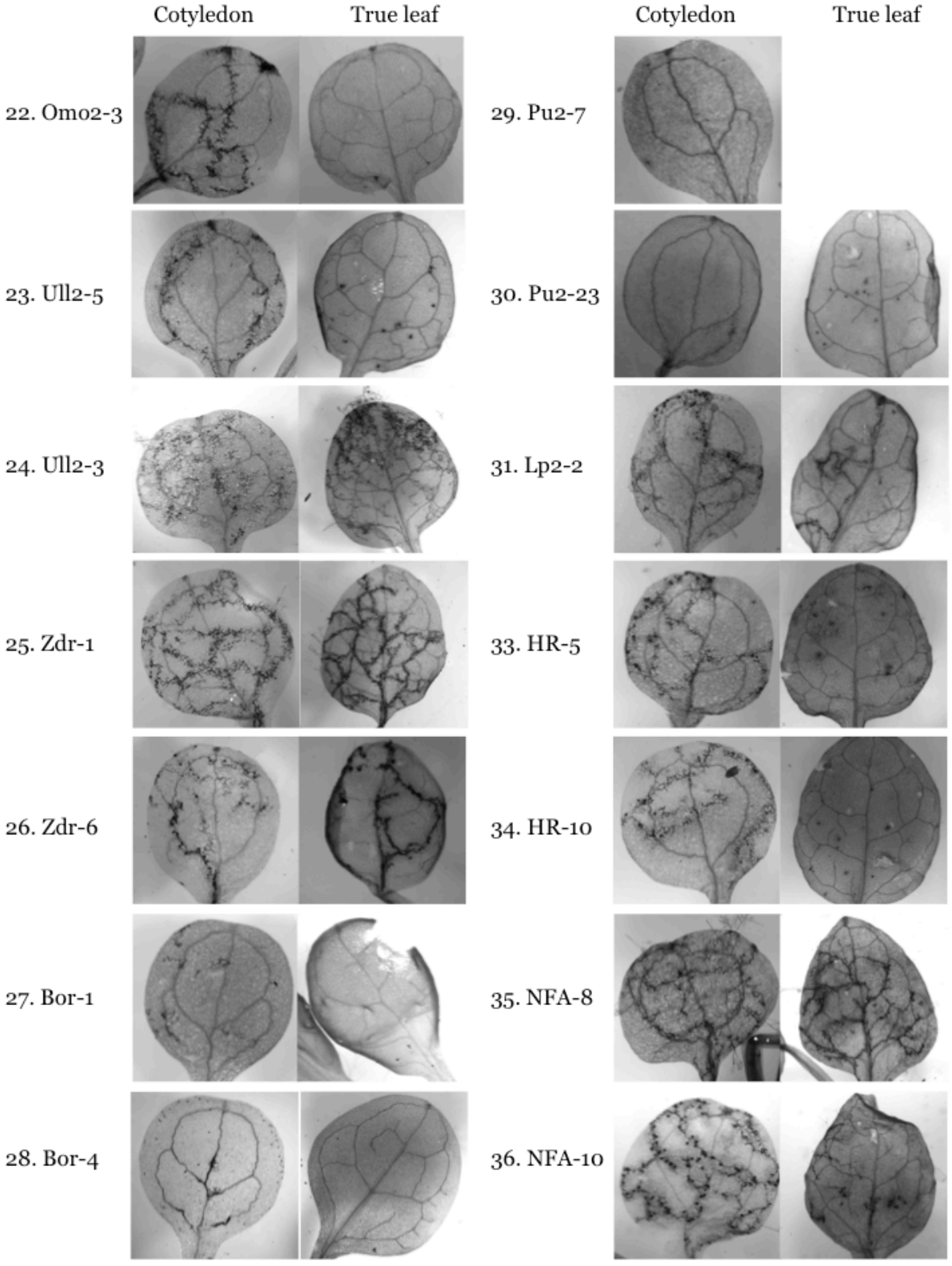
Appendix. Hpa Emco5 (cont'd)



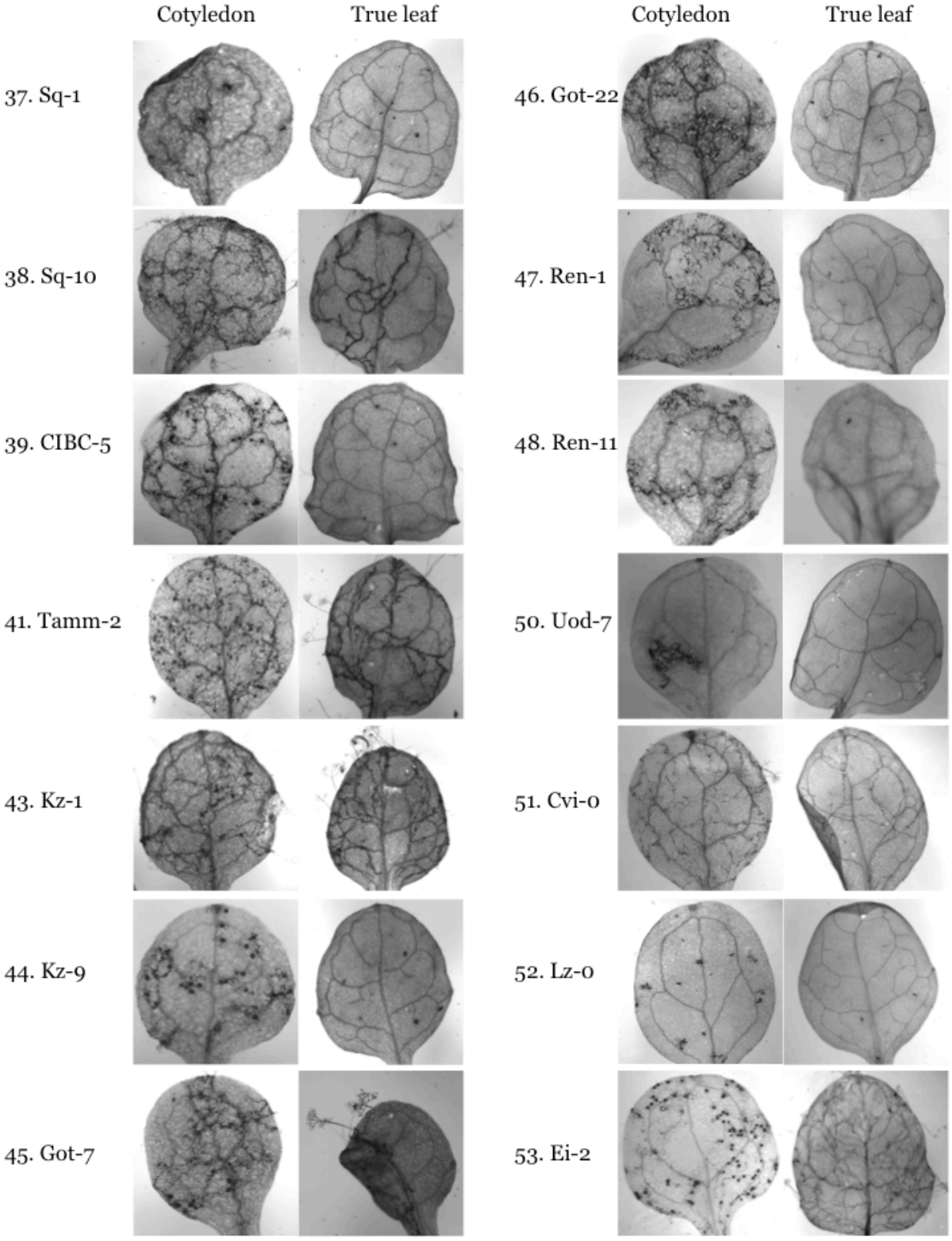
Appendix. Hpa Emwa1



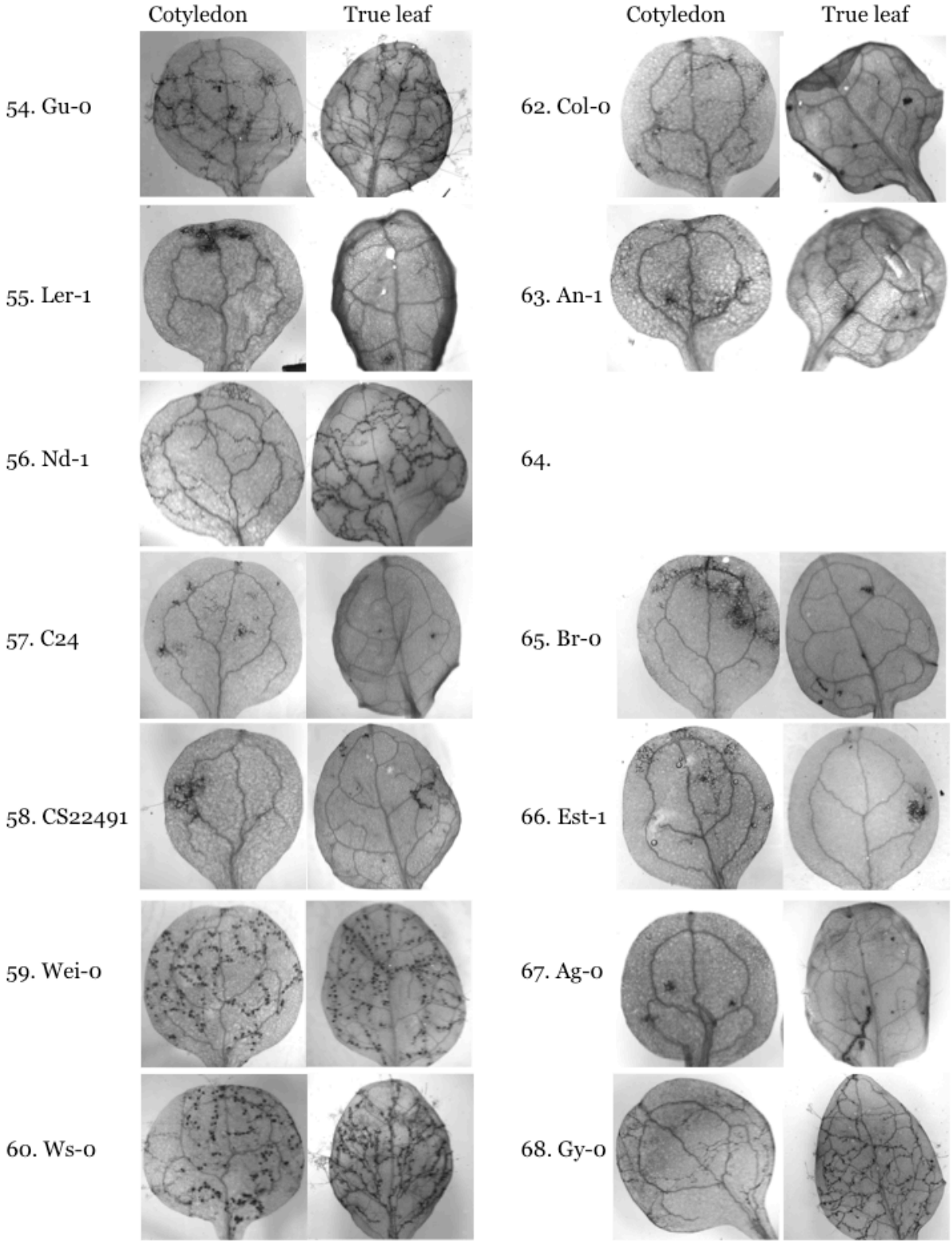
Appendix. Hpa Emwa1 (cont'd)



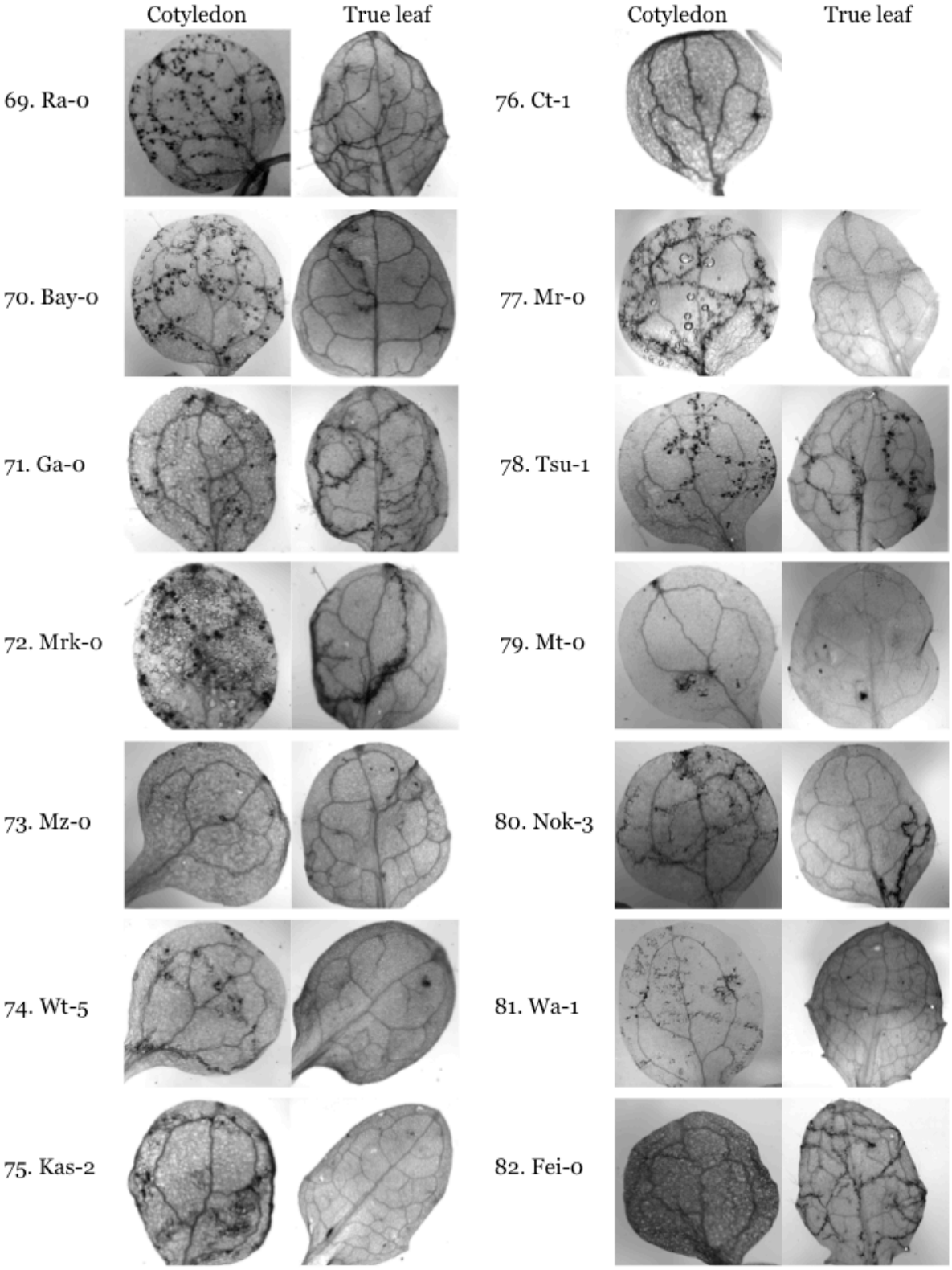
Appendix. Hpa Emwa1 (cont'd)



Appendix. Hpa Emwa1 (cont'd)



Appendix. Hpa Emwa1 (cont'd)



Appendix. Hpa Emwa1 (cont'd)